Performance of Practical Quantum Oblivious Key Distribution

Mariano Lemus,[1,*] Peter Schiansky,[2] Manuel Goulão,[3,4] Mathieu Bozzio,[2] David Elkouss,[4,5] Nikola Paunković,[6,7] Paulo Mateus,[6,7] and Philip Walther[2]

[1] *Instituto de Telecomunicações, Optical Quantum Communications Group, 3810-193 Aveiro, Portugal*

[2] *Vienna Center for Quantum Science and Technology (VCQ), Faculty of Physics, University of Vienna, Boltzmanngasse 5, Vienna A-1090, Austria*

[3] *INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal*

[4] *Networked Quantum Devices Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan*

[5] *QuTech, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, Netherlands*

[6] *Instituto de Telecomunicações, Security and Quantum Information Group, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal*

[7] *Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal*

(Received 5 March 2025; revised 4 August 2025; accepted 21 August 2025; published 10 October 2025)

Motivated by the applications of secure multi-party computation as a privacy-protecting data analysis tool, and identifying oblivious transfer as one of its main practical enablers, we propose a practical realization of randomized quantum oblivious transfer. By using only symmetric cryptography primitives to implement commitments, we construct computationally secure randomized oblivious transfer without the need for public-key cryptography or assumptions imposing limitations on the adversarial devices. We show that the protocol is secure under an indistinguishability-based notion of security and demonstrate an experimental implementation to test its real-world performance. Its security and performance are then compared to both quantum and classical alternatives, showing potential advantages over existing solutions based on the noisy storage model and public-key cryptography.

DOI: 10.1103/j8mt-c124

I. INTRODUCTION

A. State of the art

Cryptography is a critical tool for data privacy, a task deeply rooted in the functioning of today's digitalized world. Whether it is in terms of secure communication over the Internet or secure data access through authentication, finding ways of protecting sensitive data is of utmost importance. The one-time pad encryption scheme allows communication with perfect secrecy [1], at the cost of requiring the exchange of single-use secret (random) keys of the size of the communicated messages. Distribution of secret keys, therefore, is considered one of the most important tasks in cryptography. Modern cryptography relies heavily on conjectures about the computational hardness of certain mathematical problems to design solutions for the key distribution problem. However, as quantum computers threaten to make most of the currently used cryptography techniques obsolete [2], better solutions for data protection are needed. This transition towards quantum-resistant solutions becomes particularly crucial when it comes to protecting data associated with the government, finance and health sectors, being already susceptible to *intercept-now-decrypt-later* attacks. Cryptography solutions secure in a post-quantum world, where large-scale quantum computers will be commercially available, have been explored in two directions. Classical cryptography based solutions, also referred as post-quantum cryptography [3–5], involve using a family of mathematical problems that are conjectured to be resilient to quantum computing attacks. On the other hand, quantum cryptography-based solutions [6] using the laws of quantum mechanics can offer information-theoretic security, depending on the physical properties of quantum systems rather than computational hardness assumptions. Quantum key distribution (QKD) [7] is the most well studied and developed of these

*Contact author: marianojlemush@av.it.pt

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

quantum solutions, while other works beyond QKD have been proposed [8].

It is noteworthy that secure communication is not the only cryptographic task where the private data of end users may be exposed to an adversary. Cryptography beyond secure communication and key distribution includes zero-knowledge proofs, secret sharing, contract signing, bit commitment (BC), e-voting, secure data mining, etc. [9]. A huge class of such problems can be cast as multi-party computation (MPC), where distrustful parties can benefit from a joint collaborative computation on their private inputs. It requires parties' individual inputs to remain hidden from each other during the computation, among other security guarantees such as correctness, fairness, etc. [10]. Secure MPC is a powerful cryptographic tool with a vast range of applications as it allows collaborative work with private data. Generic MPC protocols work by expressing the function to evaluate as an arithmetic or Boolean circuit and then securely evaluating the individual gates. These protocols are based on one of two main fundamental primitives [11–14]: *oblivious transfer* (OT) and *homomorphic encryption*, the former of which is the focus of this work.

A 1-out-of-2 OT [15], is the task of sending two messages, such that the receiver can choose only one message to receive, while the sender remains oblivious to this choice. The original protocol, now called all-or-nothing OT, was proposed by Rabin in 1981 [16], where a single message is sent and the receiver obtains it with 1/2 probability. The two flavors of OT were later shown to be equivalent [17]. Notably, it has been shown that it is possible to implement secure MPC using only OT as a building block [18,19]. Relevant to our work is a variation of OT called random oblivious transfer (ROT), which is similar to 1-out-of-2 OT, except that both the sent messages and the receiver's choice are randomly chosen during the execution of the protocol. This can be seen as analogous to the key distribution task, in which both parties receive a random message (the key) as output. By appropriately encrypting messages using the outputs of a ROT protocol as a shared resource, it is possible to efficiently perform 1-out-of-2 OT. As an important consequence, parties expecting to engage in MPC in the future can execute many instances of ROT in advance and save the respective outputs as keys to be later used as a resource to perform fast OTs during an MPC protocol [20]. Because of this, we can think of ROT as a basic primitive for secure MPC.

In the context of quantum cryptography, OT is remarkable because, unlike classically, there exists a reduction from OT to commitment schemes [21]. This result is somewhat undermined by the existence of several theorems regarding the impossibility of unconditionally secure commitments both in classical [22] and quantum [23,24] cryptography, and it was further proven impossible in the more general abstract cryptography framework [25]. These results, in turn, imply that unconditionally secure OT itself is impossible. In light of this, approaches with different technological or physical constraints on the adversarial power have been proposed. Practical solutions based on hardware limitations, such as bounded and noisy storage [26–29], have the disadvantage that the performance of such protocols decreases as technology improves.

Computationally secure classical protocols have also been proposed [30–33], which work under the assumptions of post-quantum public-key cryptography. Alternatively, we can take advantage of quantum reduction from OT to commitments by implementing commitment schemes using (nontrapdoor) one-way functions (OWFs), such as Hash functions [34] and pseudorandom generators [35], which allows us to construct OT from symmetric cryptography primitives. The existence of general OWFs is a weaker assumption than public-key cryptography [36,37], which requires the existence of the more restrictive *trapdoor* OWFs. This difference is significant, as the latter are defined over mathematically rich structures, such as elliptic curves and lattices, and the computational hardness of the associated problems is less understood than that of their private-key counterparts. For an in-depth study of the relation between OT and OWFs see Ref. [38].

Having established that there is a theoretical merit in using computationally secure quantum protocols to implement secure MPC, it is also important to understand how practical quantum protocols compare with currently used classical solutions in security, computational and communication complexity, and practical speed in current setups. This work focuses in studying the performance of a practical quantum ROT protocol and its potential advantages compared to currently used classical solutions for OT during MPC.

The idea of using quantum conjugate coding and commitments for oblivious transfer was originally proposed by Crépeau and Kilian [17] and then refined by Bennett *et al.*, in Ref. [21] with the BBCS92 protocol (shown in Fig. 1). This construction has been extensively studied from the point of view of its theoretical security [38–43]. However, while practical security analyses and experimental implementations have been made for quantum OT in the noisy storage model [28,29], there are no works analyzing the quantum resource requirements and the resulting performance of implementing the BBCS92 protocol using existing computationally secure commitment schemes based on OWFs. Such analyses are needed to demonstrate secure experimental implementations, and provide an important step in bringing quantum OT to real-world usage.

Motivated by practical considerations, we consider Naor-style statistically binding and computationally hiding commitments, as these are well understood and efficient to implement (note that stronger commitments can be considered, such as the quantum-based commitments studied in Refs. [38,42], however, implementing those requires significantly more computational and quantum resources).

BBCS92 Quantum OT protocol

Parties: The sender Alice and the receiver Bob.

1. Alice prepares N entangled states of the form $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and, for each state prepared, sends one of the qubits to Bob.
2. Alice randomly chooses a measurement bases string $\theta^A \in \{+, \times\}^N$ and, for each $i = 1, \ldots, N$ measures her share of the i-th entangled state in the θ_i basis to obtain outcome x_i^A and the outcome string $x^A = (x_1^A, \ldots, x_N^A)$.
3. Bob uses the same process to obtain the measurement bases and outcome strings θ^B and x^B, respectively.
4. For each i, Bob commits (θ_i^B, x_i^B) to Alice.
5. Alice chooses randomly a set of indices $T \subset \{1, \ldots, N\}$ of some fixed size and sends T to Bob.
6. For each $j \in T$, Bob opens the commitments associated to (θ_j^B, x_j^B).
7. Alice checks that $x_j^A = x_j^B$ whenever $\theta_j^A = \theta_j^B$ within the test set. If the test fails Alice aborts the protocol, otherwise she sends the string θ^A to Bob.
8. Bob separates the remaining indices in two sets: I_0 - the indices where Bob's measurement bases match Alice's, and I_1 - the set of indices where their bases do not match. Then, he samples randomly c and sends the ordered pair $(I_c, I_{\bar{c}})$ to Alice.
9. Alice defines the strings $\mathbf{x}_c^A, \mathbf{x}_{\bar{c}}^A$ using the indices in the respective sets $(I_c, I_{\bar{c}})$. Then, she samples randomly a function f from a universal hash family, sends f to Bob and outputs $m_c = f(\mathbf{x}_0)$ and $m_1 = f(\mathbf{x}_{\bar{c}})$ to Bob.
10. Similarly, Bob defines the string $\mathbf{x}^B$ from the set I_0 and outputs $m_c = f(\mathbf{x}^B)$ and c.

FIG. 1. Quantum oblivious transfer protocol based on commitments.

The contributions of this work can be summarized as follows:

We introduce the definition for a quantum ROT protocol, satisfying a strong indistinguishability-based security notion equivalent to the one presented in Ref. [44], which generalizes the security of classical ROT protocols. We present a protocol that realizes said quantum ROT based on the BBCS92 construction. The protocol uses a weakly interactive string commitment scheme, which is statistically binding and computationally hiding, and can be implemented in practice using current QKD setups.

We present a formal finite-key security proof of the proposed protocol accounting for noisy quantum channels assuming only the existence of quantum-secure OWFs, together with security bounds as functions of the protocol's parameters. We also present calculations for the maximum usable channel error, as well as for the key rate as a function of the number of shared signals per instance of the protocol. Additionally, we study the composability properties of said protocol. In particular, we show that there is a family of weakly interactive commitments, which, when used in the quantum OT protocol, result in universally composable quantum OT in the classical access random oracle model. We experimentally demonstrate our protocol using current technology with a setup based on polarization-entangled photons. We also present a security analysis, which accounts for potential implementation-specific attacks and how they can be circumvented using an appropriate reporting strategy. Finally, we compare our performance results with the performance of current ROT solutions and point out the advantages and disadvantages of using quantum ROT in the context of MPC.

II. QUANTUM RANDOM OBLIVIOUS TRANSFER

In this work, the concept of indistinguishability will be often used to compare the state of systems in a "real" run of the protocol versus another "ideal" desired state. These relations are defined over families of quantum states parametrized by the security parameter of the respective protocol. Hence, indistinguishability relations are statements on the asymptotic behavior of the protocol as the security parameter is increased. For formal definitions of both statistical and computational indistinguishability see Appendix A.

When talking about two indistinguishable families $\{\rho_1^{(k)}\}$ and $\{\rho_2^{(k)}\}$, if the parameter k is implicit, we will just refer to them as ρ_1 and ρ_2 and use the following notation to denote indistinguishability:

$$\rho_1 \approx \rho_2 \quad \text{for statistically indistinguishable;}$$

$$\rho_1 \approx^{(c)} \rho_2 \quad \text{for computationally indistinguishable.}$$

Additionally, in this work we consider protocols that can abort if certain conditions are satisfied. Mathematically, it is useful to consider the state of the aborted protocol as the zero operator. This means that events that trigger the protocol to abort are described as *trace-decreasing* operations, and hence, the operator representing the associated system at the end of the protocol is, in general, not normalized. The probability of the protocol finishing successfully is given then by the trace of the final state of the output registers. Note that the above definitions of indistinguishability can be naturally extended to non-normalized operators since the outcomes of a quantum program can always be represented by the outcomes of a positive operator-valued measure (POVM) $\{F_i\}$, whose probabilities are given by $\mathrm{Tr}[F_i\rho]$, which is a well-defined quantity even for non-normalized ρ.

Definition 1 (Quantum random oblivious transfer). An n-bit quantum random oblivious transfer with security parameter k is a protocol, without external inputs, between two parties S (the sender) and R (the receiver), which, upon finishing, outputs the joint quantum state ρ_{M_0,M_1,C,M_C} satisfying the following:

(1) (Correctness) The final state of the outputs when the protocol is run with both honest parties satisfies

$$\rho_{M_0,M_1,C,M_C} \approx \frac{p_{\text{succ}}}{2^{(2n+1)}} \sum_{\substack{m_0,m_1\in\{0,1\}^n \\ c\in\{0,1\}}} \left(|m_0\rangle\langle m_0|_{M_0}|m_1\rangle\langle m_1|_{M_1}|c\rangle\langle c|_C|m_c\rangle\langle m_c|_{M_C}\right), \tag{1}$$

where $p_{\text{succ}} = \text{Tr}[\rho_{M_0,M_1,C,M_C}]$ is the probability of the protocol finishing successfully.

(2) (Security against dishonest sender) Let H_S be the Hilbert space associated to all of the sender's memory registers. For the final state after running the protocol with an honest receiver it holds that

$$\rho_{S,C} \approx \rho_S \otimes \mathsf{U}_C. \tag{2}$$

(3) (Security against dishonest receiver) Let H_R be the Hilbert space associated to all of the receiver's memory registers. For the final state after running the protocol with an honest sender, there exists a binary probability distribution given by (p_0, p_1) such that

$$\rho_{R,M_0,M_1} \approx \sum_b \left(p_b \rho^b_{R,M_{\bar{b}}} \otimes \mathsf{U}_{M_b}\right). \tag{3}$$

The above properties define statistical security for each feature of the ROT protocol. If any of them holds for the case of a dishonest party being limited to efficient quantum operations and the notion of computational indistinguishability $\approx^{(c)}$ instead, we say that the ROT protocol is computationally secure in the respective sense.

We expect the outputs m_0, m_1, c to be uniformly distributed and the receiver always obtaining the correct corresponding m_c. The first property is typically called *correctness* and it states that, when both parties follow the protocol, the probability of it not aborting *and* having incorrect outputs is negligible in the security parameter. The probability p_{succ} of the protocol finishing appears explicitly in this expression as the success of quantum protocols often depends on external conditions, most notably the noise in the quantum communication channels. For any specific value of p_{succ} and any $\varepsilon^r \leq 1 - p_{\text{succ}}$ we say that, under the associated external conditions, the protocol is $\varepsilon^{(r)}$ *robust*.

The second property, called *security against dishonest sender*, states that regardless of how much the sender deviates from the protocol, their final quantum state (which includes all the information accessible to them) is uncorrelated to the uniformly distributed value of the receiver's choice bit c. Analogously, the third property, called *security against dishonest receiver*, states that even for a receiver running an arbitrary program, by the end of the protocol there is always at least one of the two strings m_0, m_1 that is completely unknown to them (denoted by m_b).

A. Additional schemes

In this section, we define the subroutines used inside of our main protocol. We start by defining a weakly interactive commitment scheme, which gets its name from the fact that the verifier publishes a single random message at the start, which defines the operations that the committer performs.

Definition 2 (String commitment scheme). Let $k, n \in \mathbb{N}$. A weakly interactive n-bit string commitment scheme with security parameter k is a family of efficient (in n, as well as in k) programs $\texttt{com}, \texttt{open}, \texttt{ver}$

$$\begin{aligned} \texttt{com} &: \{0,1\}^n \times \{0,1\}^{n_s(k)} \times \{0,1\}^{n_r(k)} \to \{0,1\}^{n_c(k)}; \\ \texttt{open} &: \{0,1\}^n \times \{0,1\}^{n_s(k)} \to \{0,1\}^{n_o(k)}; \\ \texttt{ver} &: \{0,1\}^{n_c(k)} \times \{0,1\}^{n_o(k)} \times \{0,1\}^{n_r(k)} \to \{0,1\}^n \cup \{\perp\}, \end{aligned} \tag{4}$$

such that

(1) (correctness) $\texttt{ver}(\texttt{com}(m,s,r),\texttt{open}(m,s),r) = m$ for all $m \in \{0,1\}^n$, $s \in \{0,1\}^{n_s}$, and $r \in \{0,1\}^{n_r}$.
(2) (hiding property) For all $m_1, m_2 \in \{0,1\}^n$ and $r \in \{0,1\}^{n_r}$ the distributions for $\texttt{com}(m_1,s_1,r)$ and $\texttt{com}(m_2,s_2,r)$ are computationally (or statistically) indistinguishable in k whenever s_1, s_2 are sampled uniformly.
(3) (binding property) For uniformly sampled r, the probability $\varepsilon_{\text{bind}}(k)$ that there exists a tuple $(\text{com}, \text{open}_1, \text{open}_2)$ such that $\texttt{ver}(\text{com}, \text{open}_{1/2}, r) \neq \perp$ and

$$\texttt{ver}(\text{com}, \text{open}_1, r) \neq \texttt{ver}(\text{com}, \text{open}_2, r), \quad (5)$$

is negligible in k.

Weakly interactive string commitment schemes can be implemented using common cryptographic primitives like hash functions or pseudorandom generators. Most notably, Naor's commitment protocol [35] provides a black-box construction of weakly interactive commitments from OWFs.

Definition 3 (Verifiable information reconciliation scheme). Let $\mathcal{C} \subseteq \{0,1\}^n \times \{0,1\}^n$. A verifiable one-way information reconciliation (IR) scheme with security parameter k and leak ℓ for $\mathcal{C}$ is a pair of efficient programs $(\texttt{syn}, \texttt{dec})$ with

$$\begin{aligned} &\texttt{syn} : \{0,1\}^n \to \{0,1\}^\ell, \\ &\texttt{dec} : \{0,1\}^\ell \times \{0,1\}^n \to \{0,1\}^n \cup \{\perp\}, \end{aligned} \quad (6)$$

such that,

(1) (correctness) Whenever $(x,y) \in \mathcal{C}$ it holds that $\texttt{dec}(\texttt{syn}(x), y) = x$ except with negligible probability in k.
(2) (verifiability) For any $(x,y) \in \{0,1\}^n \times \{0,1\}^n$ it holds that either $\texttt{dec}(\texttt{syn}(x), y) = x$ or $\texttt{dec}(\texttt{syn}(x), y) = \perp$, except with negligible probability $\varepsilon_{\text{IR}}(k)$.

Due to Shannon's noisy-channel coding theorem, the size of the leak ℓ for any IR scheme over a discrete memoryless channel is lower bounded by $h(p)$, where p represents the bit-error probability, and $h(\cdot)$ denotes the binary entropy function. For concrete IR schemes, we can usually describe their efficiency using the ratio between the scheme's leak and the theoretical optimal: $f = \ell/h(p)$.

III. THE PROTOCOL

In this section we present the protocol π_{QROT} for an n-bit quantum ROT based on the primitives described in the previous section and the use of quantum communication. The protocol's main security parameter is N_0, which corresponds to the number of quantum signals sent during the quantum phase. Additionally, it has two secondary security parameters k, k', which define the security of the underlying commitment and IR schemes, respectively.

In order to facilitate the finite-key security analysis, the description of π_{QROT} features two statistical tolerance parameters, denoted as δ_1, δ_2. The role of δ_1 is to account for the error in the estimation of the qubit error rate (QBER), while the role of δ_2 is to account for the small variations in the frequency of outcomes of 50/50 events. These parameters can be ignored (set to zero) when considering very large values of N_0.

In the following description of the protocol we use the common conjugate coding notation used in BB84-based protocols. The bit values 0, 1 the denote the computational and Hadamard bases for qubit Hilbert spaces, respectively. For added clarity, we use the superscripts A and B to respectively denote Alice and Bob. Additionally, we use variable x to denote measurement outcomes and θ to denote measurement bases [e.g., the pair (θ_i^A, x_i^A) denotes that Alice measured her ith subsystem in the θ_i^A basis and obtained x_i^A as the outcome]. We use $|\Phi^+\rangle$ to denote the Bell state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Finally, we will use the relative (or normalized) Hamming weight function $r_H : \{0,1\}^n \to [0,1]$ defined for any $x = (x_1, \ldots, x_n)$ as

$$r_H(x) = \frac{1}{n}\sum_{i=1}^{n} x_i. \quad (7)$$

Parameters:

(1) Parameter estimation sample ratio $0 < \alpha < 1$.
(2) Statistical tolerance parameters δ_1, δ_2.
(3) Maximum qubit error rate $0 \leq p_{\max} \leq 1/2$.
(4) Coincidence block size $N_0 \in \mathbb{N}$, test set size $N_{\text{test}} = \alpha N_0$, minimum check set size $N_{\text{check}} = (\frac{1}{2} - \delta_2)\alpha N_0$, and raw string block size $N_{\text{raw}} = (\frac{1}{2} - \delta_2)(1-\alpha)N_0$.
(5) Weakly interactive 2-bit string commitment scheme $(\texttt{com}, \texttt{open}, \texttt{ver})$, which is computationally hiding and statistically binding, with security parameter $k \in \mathbb{N}$ and associated string lengths n_s, n_r, n_c, n_o.
(6) Verifiable one-way information reconciliation scheme $(\texttt{syn}, \texttt{dec})$ on the set $\mathcal{C} = \{(x,y) \in \{0,1\}^{N_{\text{raw}}} \times \{0,1\}^{N_{\text{raw}}} : r_H(x \oplus y) < p_{\max} + \delta_1\}$, with security parameter $k' \in \mathbb{N}$ and leak $\ell = f \cdot h(p_{\max} + \delta_1)$.
(7) Universal hash family $\mathbf{F} = \{f_i : \{0,1\}^{N_{\text{raw}}} \to \{0,1\}^n\}_i$.

Parties: The sender Alice and the receiver Bob.

Protocol steps:

Quantum phase

(1) Alice generates the state $\bigotimes_{i=1}^{N_0} |\Phi^+\rangle_i$ and sends one qubit of each entangled pair to Bob through a

(potentially noisy) quantum channel. Then she samples the string $\theta^A \in \{0,1\}^{N_0}$ and for each $i \in I = \{1,\ldots,N_0\}$ performs a measurement in the basis θ_i^A on her qubit of $|\Phi^+\rangle_i$ to obtain the outcome string x^A.

(2) Bob samples the string $\theta^B \in \{0,1\}^{N_0}$ and for each $i \in I$ performs a measurement in the basis θ_i^B on his qubit of $|\Phi^+\rangle_i$ to obtain the outcome string x^B.

Commit and open phase

(3) Alice uniformly samples the string $r \in \{0,1\}^{n_r}$ and sends it to Bob.

(4) For each $i \in I$, Bob samples a random string $s_i \in \{0,1\}^{n_s}$, computes

$$(\mathrm{com}_i, \mathrm{open}_i) = \Big(\mathtt{com}\left((\theta_i^B, x_i^B), s_i, r\right), \mathtt{open}\left((\theta_i^B, x_i^B), s_i\right)\Big), \quad (8)$$

and sends the string $\mathrm{com} = (\mathrm{com}_i)$ to Alice.

(5) Alice randomly chooses a subset test $I_t \subset I$ of size αN_0 and sends I_t to Bob.

(6) For each $j \in I_t$, Bob sends open_j to Alice.

(7) For each $j \in I_t$, Alice checks that $\mathtt{ver}(\mathrm{com}_j, \mathrm{open}_j, r) \neq \perp$. If so, she sets $(\tilde{\theta}_j^B, \tilde{x}_j^B) = \mathtt{ver}(\mathrm{com}_j, \mathrm{open}_j, r)$. Then, Alice computes the set $I_s = \{j \in I_t | \theta_j^A = \tilde{\theta}_j^B\}$ and the quantity

$$p = r_H\left(x_{I_s}^A \oplus \tilde{x}_{I_s}^B\right), \quad (9)$$

and checks that $|I_s| \geq N_{\mathrm{check}}$ and $p \leq p_{\max}$. If any of the checks fail Alice aborts the protocol.

String separation phase

(8) Alice sends $\theta_{\bar{I}_t}^A$ to Bob.

(9) Bob constructs the set I_0 by randomly selecting N_{raw} indices $i \in \bar{I}_t$ for which $\theta_i^A = \theta_i^B$. Similarly, he constructs I_1 by randomly selecting N_{raw} indices $i \in \bar{I}_t$ for which $\theta_i^A \neq \theta_i^B$. He then samples a random bit c and sends the ordered pair $(I_c, I_{\bar{c}})$ to Alice. If Bob is not able to construct I_0 or I_1, he aborts the protocol.

Post-processing phase

(10) Alice computes the strings $\left(\mathtt{syn}(x_{I_c}^A), \mathtt{syn}(x_{I_{\bar{c}}}^A)\right)$ and sends the result to Bob.

(11) Bob computes $\mathtt{dec}\left(x_{I_0}^B, \mathtt{syn}(x_{I_0}^A)\right) = y^B$. If $y^B = \perp$ Bob aborts the protocol.

(12) Alice randomly samples $f \in \mathbf{F}$, computes $m_0^A = f(x_{I_c}^A)$ and $m_1^A = f(x_{I_{\bar{c}}}^A)$, sends the description of f to Bob and outputs (m_0^A, m_1^A).

(13) Bob computes $m^B = f(y^B)$ and outputs (m^B, c).

A. Security and performance of the main protocol

We start by stating the main theorem regarding security of the proposed π_{QROT} protocol.

Theorem 1 (Security of π_{QROT}). The protocol π_{QROT} is a statistically correct, computationally secure against dishonest sender, and statistically secure against dishonest receiver n-bit ROT protocol.

A high-level proof of Theorem 1, including the derivation of the security bounds from Lemmas 1 and 2 can be found in Sec. IV and further details can be found in Appendix B. The security of π_{QROT} is given by its main security parameter N_0, as well as the security parameters of the underlying commitment and IR schemes k and k', respectively. These values can be computed for the statistical security features of the protocol and are given by the following lemmas:

Lemma 1 (Correctness). The outputs of π_{QROT} when run by honest sender and receiver satisfy

$$\rho_{M_0,M_1,C,M_C} \approx_\varepsilon \frac{p_{\mathrm{succ}}}{2^{(2n+1)}} \sum_{\substack{m_0,m_1\in\{0,1\}^n \\ c\in\{0,1\}}} \left(|m_0\rangle\langle m_0|_{M_0}|m_1\rangle\langle m_1|_{M_1}|c\rangle\langle c|_C|m_c\rangle\langle m_c|_{M_C}\right), \quad (10)$$

with

$$\varepsilon = 2^{-\frac{1}{2}(N_{\mathrm{raw}}-n)} + 2\varepsilon_{\mathrm{IR}}(k'), \quad (11)$$

where $\varepsilon_{\mathrm{IR}}$ is a negligible function given by the security of the underlying IR scheme.

Lemma 2 (Security against dishonest receiver). For the final state after running the protocol of π_{QROT} with an honest sender, there exists a binary probability distribution given by (p_0, p_1) such that

$$\rho_{R,M_0,M_1} \approx_{\varepsilon'} \sum_b \left(p_b \rho_{R,M_{\bar{b}}}^b \otimes \mathsf{U}_{M_b}\right), \quad (12)$$

with

$$\varepsilon' = \sqrt{2}\left(e^{-\frac{1}{2}(1-\alpha)^2 N_{\text{test}}\delta_1^2} + e^{-\frac{1}{2}N_{\text{check}}\delta_1^2}\right)^{\frac{1}{2}} + e^{-D_{\text{KL}}\left(\frac{1}{2}-\delta_2|\frac{1}{2}\right)(1-\alpha)N_0} + \varepsilon_{\text{bind}}(k) + \frac{1}{2}\cdot 2^{\frac{1}{2}\left(n-N_{\text{raw}}\left(\frac{1}{2}-\frac{2\delta_2}{1-2\delta_2}-h\left(\frac{p_{\max}+\delta_1}{\frac{1}{2}-\delta_2}\right)-f\cdot h(p_{\max}+\delta_1)\right)\right)}. \tag{13}$$

where $\mathcal{H}_R$ denotes the Hilbert space associated to all of the receiver's memory registers and $\varepsilon_{\text{bind}}$ is a negligible function given by the security of the underlying commitment scheme.

We can use these results to find the minimum requirements, both in terms of channel losses and number of shared entangled qubits, necessary to securely realize ROT for a given security level. We focus on the quantity

$$\varepsilon_{\max} = \varepsilon + \varepsilon'. \tag{14}$$

For the purposes of this analysis, we assume that the commitment and IR schemes, as well as their security parameters k, k', are appropriately chosen to satisfy the desired security level and we focus on the dependence of $\varepsilon_{\max}$ on the channel error rate, characterized by the parameter $p_{\max}$, and the number of quantum signals N_0. We are also interested in a quantity known as the secret key rate R_{key}. For given values of N_0, α, δ_1, δ_2, $p_{\max}$, and $\varepsilon_{\max}$, let $n_{\max}$ be the largest number for which the associated $n_{\max}$-bit ROT has at least security $\varepsilon_{\max}$, then

$$R_{\text{key}} = \frac{n_{\max}}{N_0}, \tag{15}$$

represents the ratio in which the original measurements of the shared qubits "transform" into the oblivious key. In Fig. 2 we can see the behavior of R_{key} as $p_{\max}$ increases. Note that, similarly to the case of quantum key distribution, there is a critical error p_{crit} after which R_{key} becomes negative and no secure key can be generated. The value of p_{crit} is upper bounded by ≈ 0.028, which is achieved when we set $\alpha, \delta_1, \delta_2 \to 0$, and $N_0 \to \infty$.

Another important aspect to analyze is the relation between R_{key} and N_0, which is shown in Fig. 3. Fixing the $\alpha, \delta_1, \delta_2, p_{\max}$, there is a clearly marked phase transition-like behavior in which, for each $\varepsilon_{\max}$, there is a critical value of $N_0 = N_{\text{crit}}$ before which $R_{\text{key}} = 0$, and after which it quickly reaches its maximum value. This result comes from the fact that the parameter estimation requires relatively big sample sizes to reach high confidence. It shows that, even for small n, there is a minimum amount of entangled qubits needed to be shared. In some cases, for instance, generating a 1-bit oblivious key or a 128-bit one may have similar costs in terms of quantum communication. Because the use of resources of the protocol scales with N_0, the parameters $\alpha, \delta_1, \delta_2$ should be chosen such that N_{crit} is the smallest. Figure 4 exemplifies the dependency of N_{crit} on $\varepsilon_{\max}$.

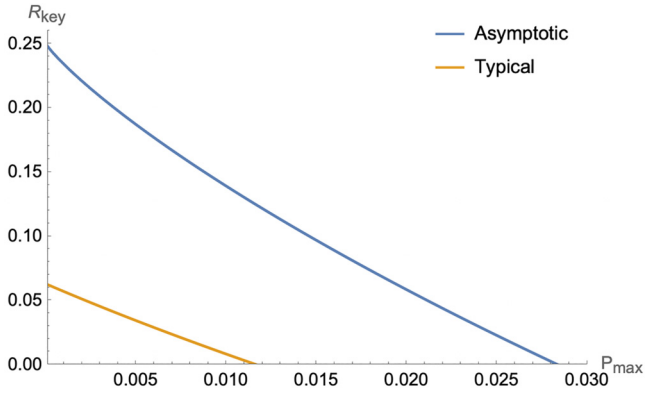


FIG. 2. Maximum key rate output n/N_0 vs error rate $p_{\max}$. The blue line represents the upper bound for the key rate, when $N_0 \to \infty$, $\alpha, \delta_1, \delta_2$ are taken to be 0 and $f = 1$. The orange line represents a more typical case with $\alpha = 0.35$, $\delta_1 = 0.01, \delta_2 = 0.025$, and $f = 1.2$.

B. Experimental implementation performance

An experiment was implemented to test the performance of the π_{QROT} protocol with contemporary technology. Data was acquired using a picosecond pulsed photon source in a Sagnac configuration [45], producing wavelength degenerate, polarization-entangled photons at 1550 nm. In this setup, entangled photons were produced via spontaneous parametric down-conversion (SPDC) by applying a laser pump beam into a 30-mm-long periodically poled potassium titanyl phosphate (ppKTP) crystal. The photon pairs

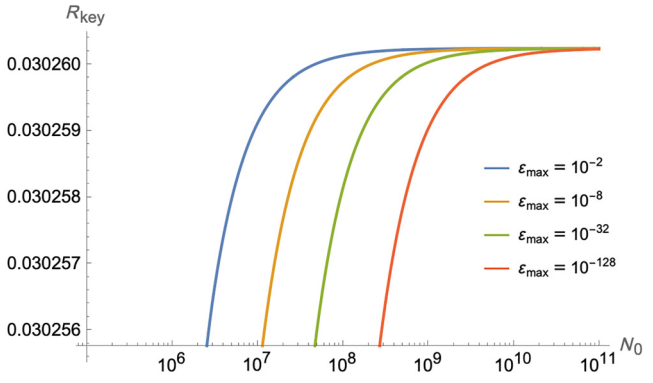


FIG. 3. Maximum key rate behavior as a function of N_0 for different security levels. Parameter values used are $\alpha = 0.35; \delta_1 = 9.20 \times 10^{-3}; \delta_2 = 3.00 \times 10^{-3}; p_{\max} = 0.01; f = 1.2$.

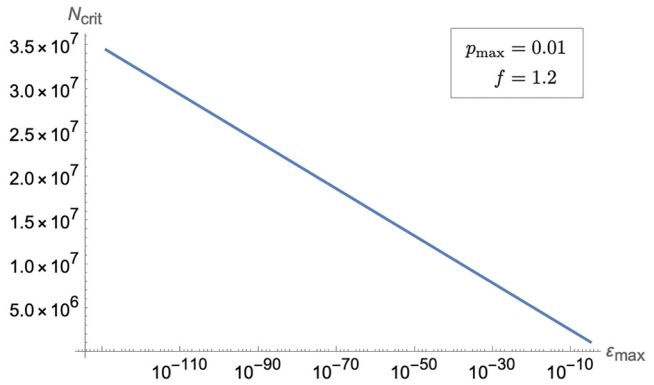


FIG. 4. Critical value N_{crit} of the number of shared qubits needed to obtain positive key rates as a function of the security level. The values of N_{crit} were computed using the parameters $\alpha, \delta_1, \delta_2$ that minimize the value of N_{crit} for each ε_{max}.

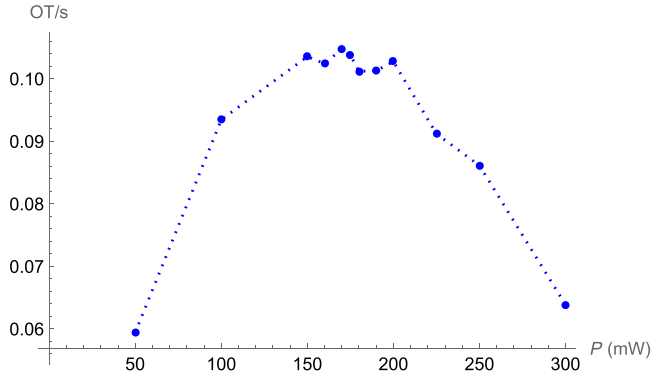


FIG. 5. Maximum potential ROT rates as a function of the pump power P for $\varepsilon_{max} = 10^{-7}$. We see that the best performance is obtained at a laser pump power of $P = 170$ mW, corresponding to a coincidence rate close to 2.45 kHz. The uncertainty on the power measurement (x axis) along with the error bars resulting from the Poissonian noise on the coincidence counts (used to calculate y values) are negligible with respect to the current plot scale.

were split using a half-wave plate (HWP) and a polarizing beam splitter (PBS), and then sent to each party where they are detected using superconducting nanowire single-photon detectors.

To test the OT speed of this implementation, different values for the power P of the laser pump were tested, as well as the use of multiplexing. As the P increases, the amount of coincidences detected per second R_c increases, but the fidelity of the produced entangled pairs decreases, resulting in larger values for qubit error rate, which is represented by the protocol parameter p_{max}. The number of maximum potential OT instances per second is computed as

$$R_{OT} = \frac{R_c}{N_{crit}}, \tag{16}$$

where N_{crit} is computed using the optimal values of $\alpha, \delta_1, \delta_2$ for the respective error rate p_{max} and undetected multi-photon rate p_{multi} associated to P, assuming perfectly efficient information reconciliation, $f = 1$ (see Sec. V for the details on the implementation and its security). As seen in Figure 5, for this implementation, the additional coincidence rate gained by increasing P is not enough to compensate for the induced increased error. This result is not immediately obvious, as N_{crit} does not depend explicitly on p_{max}. The decrease in performance comes from the fact that increasing p_{max} limits the values that δ_1 can have while maintaining positive key rates. This restriction on the values of δ_1 ultimately results in an increase in N_{crit} and therefore, a reduction on R_{OT}.

Table I shows an example of the performance of the protocol in a real-world implementation using the data from the experimental setup. For the commit and open phase, the weakly interactive string commitment protocol introduced in Ref. [35] was implemented using the BLAKE3 hash function algorithm as a one-way function. For the post-processing phase, a low-density parity-check (LDPC) code was used for IR, and random binary matrices were used to implement the universal hash family for privacy amplification. We evaluated the performance by the number of 128-bit ROT instances able to be completed per second. (It is worth noting that, using a Mac mini M1 2020 16 GB computer, the post-processing throughput was enough to handle all the data from the experiment, the bottleneck being the quantum signal generation rate.) Detector measurement data is publicly available in Ref. [46].

IV. SECURITY ANALYSIS

In this section, we prove the main security result, which relates the overall security of the protocol as a function of

TABLE I. Table of protocols parameters and the resulting performance. The values of N_0 and δ_1 and the laser pump power were optimized to yield the highest ROT rate for an LDPC code with efficiency $f = 1.61$.

Parameter	Symbol	Value
Message size (bits)	n	128
Security level	ε_{max}	1.91×10^{-8}
Cost in quantum signals	N_0	5.86×10^6
Max allowed QBER	p_{max}	1.14%
Testing set ratio	α	0.35
Statistical parameter 1	δ_1	9.00×10^{-3}
Statistical parameter 2	δ_2	3×10^{-3}
IR verifiability security	ε_{IR}	2^{-32}
Commitment binding security	ε_{bind}	2^{-32}
Efficiency of IR	f	1.64
Max allowed multi-photon rate	p_{multi}	3.67×10^{-3}
	ROT rate	0.023 ROT/s

its parameters N_0, α, δ_1, and δ_2 in Theorem 1. For clarity of presentation, we have compacted some of the properties into lemmas, for which detailed proofs can be found in Appendix B. Definitions and properties of entropic quantities can be found in Appendix A

A. Correctness

In order to prove correctness we need to show that either the protocol either finishes with Alice outputting uniformly distributed messages m_0, m_1 and Bob outputting a uniformly random bit c and the corresponding message m_c, or it aborts, except with negligible probability.

Recall that we model the aborted state of the protocol as the zero operator. This way, whenever we have a mixture of states, some of which trigger aborting and some that do not, the abort operation removes the events that trigger it from the mixture, effectively reducing its trace by the probability of aborting. There are three instances where the protocol can abort: first during step (7) if the estimated qubit error rate is larger than $p_{\max}$; the second one is during step (9) if Bob does not have enough (mis)matching bases to construct the sets I_0, I_1; and finally during step (11) if the IR verification fails. The probability of aborting in steps (7) and (11) depends on the particular transformation that the states undergo when being sent from Alice's to Bob's laboratory, about which we make no assumptions. We can group these three abort events and denote by p_{abort} the probability of the protocol aborting by the end of step (11). The state at this point can be written as $(1-p_{\text{abort}})\rho^{\top}$, where $\rho^{\top}$ represents the normalized state conditioned that the protocol has not aborted by this point. As Lemma 3 states, the verifiability property of the information reconciliation scheme guarantees that the states that "survive" past Step (11) have the property that Bob's corrected string y^B is the same as Alice's outcome string $x^A_{I_0}$, which is uniformly distributed.

Lemma 3. Let $X^A_{I_0}, X^A_{I_1}, C, Y^B$ denote the systems holding the information of the respective values $x^A_{I_0}, x^A_{I_1}, c$, and y^B of π_{QROT}. Denote by $\rho^{\top}$ the parties' joint state at the end of step (11) conditioned that Bob constructed the sets (I_0, I_1) during step (9) and the protocol has not aborted. Assume both parties follow the steps of the protocol, then

$$\rho^{\top}_{X^A_{I_0},X^A_{I_1},C,Y^B} \approx_{\varepsilon_{\text{IR}}(k')} \tilde{\rho}^{\top}_{X^A_{I_0},X^A_{I_1},C,Y^B}, \tag{17}$$

where $\varepsilon_{\text{IR}}(k')$ is a negligible function given by the security of the underlying information reconciliation scheme, k' its associated security parameter, and

$$\tilde{\rho}^{\top}_{X^A_{I_0},X^A_{I_1},C,Y^B} = \frac{1}{2^{(2N_{\text{raw}}+1)}} \sum_{\substack{x_{I_0},x_{I_1}\\ c}} |x_{I_0}\rangle\langle x_{I_0}|_{X^A_{I_0}} |x_{I_1}\rangle\langle x_{I_1}|_{X^A_{I_1}} |x_{I_0}\rangle\langle x_{I_0}|_{Y^B} |c\rangle\langle c|_C. \tag{18}$$

During step (12) universal hashing is used in both $x^A_{I_0}$ and $x^A_{I_1}$ to obtain m_c and $m_{\bar{c}}$. Because Eq. (18) describes a state for which the $X^A_{I_0}, X^A_{I_1}$, and C subsystems are independent and uniformly distributed, it follows from Lemma 3 that

$$\begin{aligned} H^{\varepsilon_{\text{IR}}(k')}_{\min}(X^A_{I_0}Y^B|X^A_{I_1}C)_{\rho^{\top}} &= H^{\varepsilon_{\text{IR}}(k')}_{\min}(X^A_{I_1}|X^A_{I_0}Y^BC)_{\rho^{\top}} \\ &= N_{\text{raw}}. \end{aligned} \tag{19}$$

Finally, using the quantum leftover hash Lemma A4 twice (once for m_0 and m_1) with the corresponding entropy terms given by Eq. (19), together with Lemma A1 (1), we conclude that the state $\rho^{(\text{out})}_{M_0,M_1,C,M_C}$ of the output systems after the post-processing phase satisfies (substituting $p_{\text{succ}} = 1 - p_{\text{abort}}$)

$$\rho^{(\text{out})}_{M_0,M_1,C,M_C} \approx_{\varepsilon} \frac{p_{\text{succ}}}{2^{(2n+1)}} \sum_{\substack{m_0,m_1\in\{0,1\}^n\\ c\in\{0,1\}}} |m_0\rangle\langle m_0|_{M_0} |m_1\rangle\langle m_1|_{M_1} |c\rangle\langle c|_C |m_c\rangle\langle m_c|_{M_C}, \tag{20}$$

with

$$\varepsilon \leq 2^{-\frac{1}{2}(N_{\text{raw}}-n)} + 2\varepsilon_{\text{IR}}(k'). \tag{21}$$

B. Security against dishonest sender

For this scenario we show that, in the case of an honest Bob and Alice running an arbitrary program, the resulting state after the protocol successfully finishes satisfies Eq. (2). In other words, independently of what quantum state Alice shares at the beginning of the protocol and which operations she performs on her systems, her final state is independent of the value of c. We assume that Alice's laboratory consists of everything outside Bob's. In particular, this means that she controls the environment, which includes the transmission channels. We also assume that Alice is limited to performing efficient computations.

Let A be the system consisting of all of Alice's laboratory after step (1) of the protocol, that is, A contains her part of the shared system and every other ancillary system she may have access, but does not contain any system from Bob's laboratory, including Bob's part of the system shared in step (1). During the execution of the protocol, Alice receives external information from Bob exactly three times: the commitment information shared during step (4), the opening information open_{I_t} for the commitments associated to the test set I_t in step (6), and the information of the pair of sets $(J_0, J_1) = (I_c, I_{\bar{c}})$ during step (9). Let $\text{COM} = (\text{COM}_i)_{i=1}^{N_0}$ and $\text{OPEN} = (\text{OPEN}_i)_{i=1}^{N_0}$ be the respective systems used by Bob to store the information of the strings $\text{com} = (\text{com}_i)_{i=1}^{N_0}$ and $\text{open} = (\text{open}_i)_{i=1}^{N_0}$, and let SEP be the system holding the string separation information (J_0, J_1). We want to show that, by the end of the protocol, the state of the system $A, \text{COM}, \text{OPEN}_J, \text{SEP}, C$ satisfies

$$\rho_{A,\text{COM},\text{OPEN}_{\bar{I}_t},\text{SEP},C} \approx^{(c)} \rho_{A,\text{COM},\text{OPEN}_{\bar{I}_t},\text{SEP}} \otimes \mathsf{U}_C. \tag{22}$$

To guarantee that Alice will not be able to obtain information about the value of c during the string separation phase, it is necessary to show that Alice does not have access to the information of Bob's bases choices $\theta^B_{I_0,I_1}$ from the commitments sent by Bob during step (4) of the protocol. As shown by Lemma 4, the shared state of the parties after the commitment information is sent is computationally indistinguishable from a state where Alice's information is independent of $\theta^B_{\bar{I}_t}$.

Lemma 4. Assuming Bob follows the protocol, for any $J \subseteq I$, the state of the system $A, \text{COM}, \text{OPEN}_J, \Theta^B_{\bar{J}}$ after step (4) satisfies

$$\rho_{A,\text{COM},\text{OPEN}_J,\Theta^B_{\bar{J}}} \approx^{(c)} \rho_{A,\text{COM},\text{OPEN}_J} \otimes \mathsf{U}_{\Theta^B_{\bar{J}}}. \tag{23}$$

At step (8) of the protocol, Alice sends Bob the system $\Theta^A_{\bar{I}_t}$ intended to have the information of her measurement bases. Bob then is able to determine the indices for which $\theta^A_{\bar{I}_t}$ and $\theta^B_{\bar{I}_t}$ coincide. With this information, he randomly selects sets $I_0, I_1 \in \bar{I}_t$ of size N_{raw} for which all indices are associated with matching (for I_0) or nonmatching (for I_1) bases. Then he computes $(J_0, J_1) = (I_c, I_{\bar{c}})$, by flipping the order if the pair (I_0, I_1) depending on the value of c. Clearly, (J_0, J_1) depend on both $\theta^B_{\bar{I}_t}$ and c, but as Lemma 5 states, any correlation between (J_0, J_1), c, and Alice's information disappears if one does not have access to $\theta^B_{\bar{I}_t}$.

Lemma 5. Denote by A' the system representing Alice's laboratory at the start of step (9). Let $\mathcal{E}^{(I_t)} : \mathcal{D}(\mathcal{H}_{A',\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C}) \to \mathcal{D}(\mathcal{H}_{A',\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C,\text{SEP}})$ be the quantum operation used by Bob to compute the string separation information (J_0, J_1) during step (9) of the protocol. The resulting state after applying $\mathcal{E}^{(I_t)}$ to a product state of the form

$$\mathcal{E}^{(I_t)}(\rho_{A',\Theta^A_{\bar{I}_t}} \otimes \mathsf{U}_{\Theta^B_{\bar{I}_t}} \otimes \mathsf{U}_C) = \sigma_{A',\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C,\text{SEP}} \tag{24}$$

satisfies

$$\text{Tr}_{\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t}}\left[\sigma_{A',\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C,\text{SEP}}\right] = \sigma_{A'} \otimes \sigma_{\text{SEP}} \otimes \mathsf{U}_C. \tag{25}$$

A proof of both Lemmas 4 and 5 can be found in Appendix B 3. By setting $J = I_t$, Lemma 4 guarantees that Alice's system's state after the opening information has been sent is computationally indistinguishable from one that is completely uncorrelated with Bob's measurement basis information in $\bar{I}_t$.

Additionally, by recalling that the value of c is sampled independently of any of the considered systems, we know that the state $\rho_{A',\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C}$ before (J_0, J_1) is computed has the required product form and, from Lemma 5, we conclude that the state of all of Alice's system at this point is computationally indistinguishable from a state uncorrelated with C. Let $\mathcal{E}$ be the operation Alice performs in her system from here to the end the protocol. By using Lemma A2 (4) and grouping all of Alice's systems into S, we obtain the desired result:

$$\rho_{S,C} \approx^{(c)} \rho_S \otimes \mathsf{U}_C. \tag{26}$$

C. Security against dishonest receiver

We consider now the scenario in which Alice runs the protocol honestly and Bob runs an arbitrary program. For this analysis, note that Alice trusts her quantum state preparation and detection. We want to show that the state after finishing the protocol successfully satisfies Eq. (3). This means that the state at the end of the protocol can be described as a mixture of states where Bob's system is uncorrelated with at least one of the two strings outputted by Alice. Similarly to the dishonest sender's case, we assume that Bob's laboratory consists of everything

outside Alice's, which means that he controls the communication channels and the environment. However, we do not assume that Bob is restricted to efficient computations.

The values of Alice's output strings depend on several quantities: Alice's measurement outcomes, the choice of the I_t, J_0, J_1 subsets, and the choice of hashing function f during the post-processing phase of the protocol. From all of these, the only ones that are not made explicitly public during the protocol's execution are Alice's measurement outcomes. Instead, partial information of these outcomes is revealed at different steps of the protocol. Let $x^A_{J_0}, x^A_{J_1}$ be the sub-strings of measurement outcomes used to compute Alice's outputs m_0, m_1, respectively, and let R denote Bob's system at the end of the protocol (which includes all the systems that Alice sent during the execution of the protocol). In order to prove security we need to show that the joint state of the system $X^A_{J_0}, X^A_{J_1}, R$ can be written as a mixture of states ρ^b (with $b \in \{0,1\}$) such that the conditional min-entropy $H^{\varepsilon}_{\min}(X^A_{J_b} \mid R)_{\rho^b}$ is high enough, so that we can use the leftover hash Lemma A4 to guarantee that the outcome of the universal hashing $m_b = f(x^A_{J_b})$ is uncorrelated with R.

At the start of the protocol the parties share a completely correlated entangled system. If the parties make measurements as intended, their outcomes will be only partially correlated, but if Bob was able to postpone his measurement until after Alice's reveals her measurement bases, Bob could potentially obtain the whole information of x^A by measuring in the appropriate basis on his system. To prevent this, Bob is required to commit his measurement bases and results to Alice before knowing which set is going to be tested. Then a statistical test is performed in step (7) to estimate the correlation of Alice's measurement outcomes with with the ones that Bob committed. As Lemma 6 states, any state passing the aforementioned test is such that, regardless of how Bob defines the sets (J_0, J_1) during the string separation phase, there is a minimum of uncertainty that he has with respect to Alice's measurement outcomes. Recall that, when Alice is honest, the overall state of the protocol before step (8) will be a partially classical state, which could be written as a mixture over all of Alice's classical information. Let $\vec{\tau} = (x^A_{I_t}, \theta^A, r, \mathrm{com}, I_t, I_s, \mathrm{open}_{I_t})$ denote the *transcript of the protocol* up to step (8), and let $\rho_{X^A B}(\vec{\tau}, J_0, J_1)$ be the joint state of Alice's measurement outcomes and Bob's laboratory conditioned to $\vec{\tau}, J_0, J_1$.

Lemma 6. Assuming Alice follows the protocol, let T, SEP, B denote the systems of the protocols transcript, the strings J_0, J_1, and Bob's laboratory at the end of step (9) of the protocol, and let $\rho_{T,\mathrm{SEP},X^A,B}$ be the state of the joint system at that point. There exists a state $\tilde{\rho}_{T,\mathrm{SEP},X^A,B}$, which is classical in T and SEP such as the following:

(1) The conditioned states $\tilde{\rho}_{X^A,B}(\vec{\tau}, J_0, J_1)$ satisfy

$$H_{\min}(X^A_{J_0} \mid X^A_{J_1} B)_{\tilde{\rho}(\vec{\tau},J_0,J_1)} + H_{\min}(X^A_{J_1} \mid X^A_{J_0} B)_{\tilde{\rho}(\vec{\tau},J_0,J_1)} \geq 2N_{\mathrm{raw}} \left(\frac{1}{2} - \delta_2 - h\left(\frac{p_{\max} + \delta_1}{\frac{1}{2} - \delta_2} \right) \right), \tag{27}$$

(2) $\rho_{T,\mathrm{SEP},X^A,B} \approx_{\varepsilon} \tilde{\rho}_{T,\mathrm{SEP},X^A,B}$, with

$$\varepsilon = \left(2 \left(e^{-\frac{1}{2}\alpha(1-\alpha)^2 N_0 \delta_1^2} + e^{-\frac{1}{2}\left(\frac{1}{2}-\delta_2\right)\alpha N_0 \delta_1^2} \right) \right)^{\frac{1}{2}} + e^{-D_{\mathrm{KL}}\left(\frac{1}{2}-\delta_2 \mid \frac{1}{2}\right)(1-\alpha)N_0} + \varepsilon_{\mathrm{bind}}(k), \tag{28}$$

where $h(\cdot)$ and $D_{\mathrm{KL}}(\cdot \mid \cdot)$ denote the binary entropy and the binary relative entropy functions, respectively, and $\varepsilon_{\mathrm{bind}}(k)$ is a negligible function given by the binding property of the commitment scheme.

To reach the desired result, we will first show that a state $\tilde{\rho}_{T,\mathrm{SEP},X^A,B}$ satisfying Lemma 6 (1) also satisfies a tighter version Lemma 2, and then use Lemma 6 (2) to attain the bound for the real protocol's outcome. Since $\tilde{\rho}_{T,\mathrm{SEP},X^A,B}$ is classical in both T and SEP we can write the state of the joint system of Alice's measurement outcomes and Bob's laboratory as a mixture over all the possible transcripts at that point, that is

$$\tilde{\rho}_{X^A B} = \sum_{\vec{\tau}, J_0, J_1} P(\vec{\tau}, J_0, J_1) \tilde{\rho}_{X^A B}(\vec{\tau}, J_0, J_1), \tag{29}$$

where $P(\vec{\tau}, J_0, J_1)$ defines a probability distribution, which is dependent on Bob's behavior during the previous steps. We can now separate the $\tilde{\rho}_{X^A B}(\vec{\tau}, J_0, J_1)$ in two categories depending on which of the $x^A_{J_0}, x^A_{J_1}$ is the least correlated with Bob's system. Consider the function $b(\vec{\tau}, J_0, J_1)$ to be equal to 0 if $H^{\varepsilon}_{\min}(X^A_{J_0} \mid X^A_{J_1} B)_{\rho(\vec{\tau}, J_0, J_1)} \geq H^{\varepsilon}_{\min}(X^A_{J_1} \mid X^A_{J_0} B)_{\rho(\vec{\tau}, J_0, J_1)}$, and equal to 1 otherwise. By regrouping the terms from Eq. (29) for which the value of b is the same, we can rewrite the joint state as

$$\tilde{\rho}_{X^A B} = \sum_{b \in \{0,1\}} P_b \tilde{\rho}^b_{X^A B}, \tag{30}$$

where, from Lemma 6 and recalling that, as Lemma A3 (5) states, the min-entropy of a mixture is lower bounded by that of the term with the least min-entropy, we know that

$$H_{\min}(X^A_{J_b} \mid X^A_{J_{\bar{b}}} B)_{\tilde{\rho}^b} \geq N_{\text{raw}} \left(\frac{1}{2} - \frac{2\delta_2}{1 - 2\delta_2} - h\left(\frac{p_{\max} + \delta_1}{\frac{1}{2} - \delta_2} \right) \right). \tag{31}$$

At step (10), Alice shares with Bob the syndromes $S_0 = \mathtt{syn}(x^A_{J_0})$ and $S_1 = \mathtt{syn}(x^A_{J_1})$. Since these syndromes are completely determined by the respective sub-strings $x^A_{J_i}$, we know that

$$\begin{aligned} H_{\min}(X^A_{J_b} \mid S_b S_{\bar{b}} B) &\geq H_{\min}(X^A_{J_b} \mid S_b X^A_{J_{\bar{b}}} B) \\ &\geq H_{\min}(X^A_{J_b} \mid X^A_{J_{\bar{b}}} B) - H_{\max}(S_b) \\ &\geq N_{\text{raw}} \left(\frac{1}{2} - \frac{2\delta_2}{1 - 2\delta_2} - h\left(\frac{p_{\max} + \delta_1}{\frac{1}{2} - \delta_2} \right) - f \cdot h(p_{\max} + \delta_1) \right), \end{aligned} \tag{32}$$

where the second inequality follows from Lemma A3 (3) and (4), and the max-entropy term $H_{\max}(S_b)$ is upper bounded by the size in bits of the syndrome $\ell = N_{\text{raw}}(f \cdot h(p_{\max} + \delta_1))$. Now we can apply Eq. (32) to Lemma A4, which states that, for the outcomes M_0, M_1 of the universal hashing by Alice in step (12) and Bob's final system R it holds that

$$\tilde{\rho}^b_{R, M_0, M_1} \approx_{\varepsilon'} \tilde{\rho}^b_{R, M_{\bar{b}}} \otimes \mathsf{U}_{M_b}, \tag{33}$$

with

$$\varepsilon' = \frac{1}{2} \cdot 2^{\frac{1}{2}\left(n - N_{\text{raw}} \left(\frac{1}{2} - \frac{2\delta_2}{1 - 2\delta_2} - h\left(\frac{p_{\max} + \delta_1}{\frac{1}{2} - \delta_2} \right) - f \cdot h(p_{\max} + \delta_1) \right) \right)}. \tag{34}$$

Finally, by applying Lemma A1 (3) and (4) to Eqs. (30) and (33), and then Lemma A1 (1) to Eq. (28) we get the desired result:

$$\rho_{R, M_0, M_1} \approx_{\varepsilon} \sum_b P_b \rho^b_{R, M_{\bar{b}}} \otimes \mathsf{U}_{M_b}, \tag{35}$$

with

$$\begin{aligned} \varepsilon = \sqrt{2} \left(e^{-\frac{1}{2}(1-\alpha)^2 N_{\text{test}} \delta_1^2} + e^{-\frac{1}{2} N_{\text{check}} \delta_1^2} \right)^{\frac{1}{2}} + e^{-D_{\text{KL}}\left(\frac{1}{2} - \delta_2 \middle| \frac{1}{2} \right)(1-\alpha) N_0} + \varepsilon_{\text{bind}}(k) \\ + \frac{1}{2} \cdot 2^{\frac{1}{2}\left(n - N_{\text{raw}} \left(\frac{1}{2} - \frac{2\delta_2}{1 - 2\delta_2} - h\left(\frac{p_{\max} + \delta_1}{\frac{1}{2} - \delta_2} \right) - f \cdot h(p_{\max} + \delta_1) \right) \right)}. \end{aligned} \tag{36}$$

D. Composability considerations

Since OT protocols are mainly used as a subroutine of larger applications it is important to understand the composability properties of π_{ROT}. In general, this is done through simulation-based composability frameworks. As mentioned in Sec. I, this protocol is based on the BBCS92

construction, which has been proven secure in the quantum universal composability (UC) framework by Unruh [41] assuming access to an ideal commitment functionality. This means that we can understand the composability properties of π_{ROT} by understanding the respective properties of the underlying weakly interactive commitment protocol.

It is well known that UC commitments are impossible to realize in the plain model [22,25]. Because of this, protocols are often analyzed within a hybrid model, where the parties have access to some base external functionality. We show in Appendix IV D that there exists a family of commitment schemes that are both weakly interactive and UC-secure in the *classical access* random oracle model (ROM) [47]. This, in tandem with the aforementioned reduction of OT to commitments, results in the following theorem:

Theorem 2. There exists a family of weakly interactive commitment schemes in relation to which π_{ROT} is UC-secure in the classical access ROM.

In relation to Theorem 2, we want to emphasize that, even though limiting the access to the random oracle to be classical may seem at first strong in the context of a quantum protocol (where the parties are required access to some quantum capabilities), it has little impact in the resulting security of larger MPC protocols for which the security is analyzed in the classical setting.

Finally, we would like to stress the merits of Definition 1 by itself. In particular, this definition was studied in Refs. [48] and [44] and stated to ensure security when the protocol is executed sequentially. Furthermore, the indistinguishability properties stated in Definition 1 provide a very strong security guarantee and, because the protocol does not have external inputs and the indistinguishability relations include arbitrary external systems, these properties will still hold in any environment, which makes it relatively straightforward to analyze as part of bigger applications.

V. EXPERIMENTAL IMPLEMENTATION

A. Description of the setup

A schematic representation of the experimental setup can be seen in Fig. 6. Spontaneous parametric down-conversion (SPDC), attributed to Alice, is used to create polarization entangled photon pairs in the state $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle)$, which are coupled into optical fiber. One photon is sent through a 50:50 fiber beam splitter, probabilistically routing it to one of two polarization projection stages. There, a quarter-wave plate (QWP), a half-wave plate (HWP), and a polarizing beam splitter (PBS) are used to project the photons state onto the linear (H/V) or diagonal ($+/-$) basis, respectively. All photons

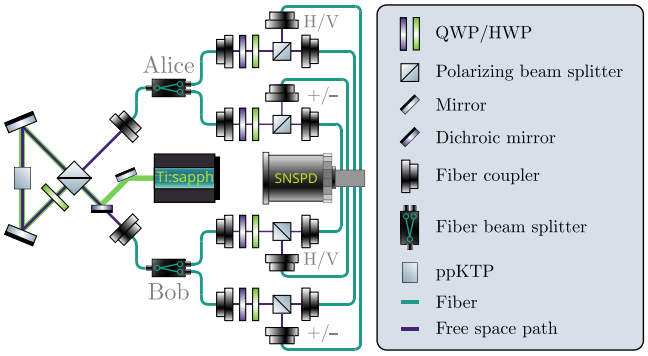


FIG. 6. Experimental setup. Polarization-entangled photon pairs are created using spontaneous parametric down-conversion. Alice's and Bob's photons are individually fiber coupled and each sent to 50:50 fiber beam splitters, which probabilistically route them to free-space polarization projection stages—one projecting onto the linear, and one onto the diagonal basis each for Bob and Alice.

are sent to superconducting nanowire single photon detectors (SNSPDs) and their arrival time is recorded using a time tagging module (TTM). The second photon of the state $|\Psi^+\rangle$, attributed to Bob, travels through an equivalent probabilistic projection setup.

Entangled photon pairs are generated using collinear type-II SPDC in a periodically poled $KTiOPO_4$ crystal with a poling period of 46.2 μm inside of a Sagnac interferometer. The pump light is produced by a pulsed Ti:sapphire laser (Coherent Mira 900HP) with a pulse width of 2.93 ps and a central wavelength of $\lambda_p = 773$ nm, creating degenerate single-photon pairs at $\lambda_s = \lambda_i = 1546$ nm. The laser's inherent pulse repetition rate of 76 MHz is doubled twice to 304 MHz using a passive temporal multiplexing scheme [49]. More precisely, for n simultaneously emitted pairs and k multiplexing stages, each doubling the repetition rate, higher-order pair production events are attenuated by a factor of $1/(2^k)^{n-1}$. In our experiment, $k = 2$, so this scheme reduces the probability of emitting a double pair ($n = 2$) by a factor of 4 compared to a source relying on the pump's inherent repetition rate, while the single-pair emission probability remains constant. Finally, about 100 m of single-mode fiber separate the experimental setup from the 1-K cryostat housing the SNSPDs with a detection efficiency of around 95% and a dark-count rate of around 300 Hz.

We note that our entanglement-based implementation presents two main technological advantages over prepare-and-measure configurations:

(1) It circumvents the need for a certified quantum random number generator or for classical pseudo-randomness that may compromise the security of the quantum phase: instead of feeding random (or

pseudorandom) sequences into the active polarization modulator of a prepare-and-measure scheme, the choice of BB84 state is performed in a passive and uniformly random way by the beam splitters present in both Alice and Bob's measurement setups (also known as "remote state preparation").

(2) In free space, it avoids the need for active polarization modulation, which imposes a strict upper limit on the protocol's repetition rate governed by the bandwidth of the Pockels cell and its high-voltage amplifier, typically achieving a few hundred kHz to a few hundred MHz [50]. By generating entangled photons that are passively projected onto one of the four BB84 states instead, our OT rate is not limited by any active prepare-and-measure encoding routine, but only by our picosecond-pulsed pump rate of around 300 MHz. With other SPDC sources reaching the GHz [51] to tens of GHz regimes [52], our passive state-preparation routine can perform even better.

B. Practical protocol

The protocol is identical to that from π_{ROT} as described in Sec. III, with the following amendments:

(1) The parties agree on an additional parameter p_{multi}—the accepted ratio of multi-photon events.

(2) During the quantum phase of the protocol, Alice may observe detection patterns that are incompatible with the emission of a single photon pair. Instead of sharing N_0 states in step (1), she continues sharing states until, after agreeing on coincidence time-tags with Bob, the parties obtain N_0 coincidences associated with single-photon events on Alice's side. Let N_{tot} be the number of coincidences obtained at this point and $N_{\mathrm{multi}} = N_{\mathrm{tot}} - N_0$. Alice computes the value

$$p'_{\mathrm{multi}} = \frac{N_{\mathrm{multi}}}{3N_{\mathrm{tot}}}, \tag{37}$$

and aborts the protocol if $p'_{\mathrm{multi}} \geq p_{\mathrm{multi}}$.

(3) Similarly to Alice, Bob may also observe multi-click patterns. While reporting its detection events he uses the following rules:

   (a) 1 click: assign the correct measured bit value and report a successful round.
   (b) 2 clicks from the same basis: assign a random bit value to the measurement result and report a successful round.
   (c) any other click pattern: report an unsuccessful round.

C. Practical security

Any photonic implementation of quantum cryptography presents experimental imperfections, which can be exploited by dishonest parties to enhance their cheating probability and violate ideal security assumptions. Important examples of such imperfections include multi-photon noise, lossy and noisy quantum channels, non-unit detection efficiency and detector dark counts.

Dishonest sender. In our experiment, threshold detectors cannot resolve the incident photon number, and unexpected click patterns can occur. For example, several of the four detectors may simultaneously click for a given round, which leads to an inconclusive measurement outcome that has to be back-reported by the honest receiver. This in turn allows a dishonest sender to gain a significant amount of information about the receiver's measurement basis choice. Adopting the reporting strategy presented above makes the protocol secure against this type of attack. For a complete analysis of both the attack and its countermeasures, see Ref. [53].

Dishonest receiver. Due to Poisson statistics in the SPDC process, emission of double pairs can occur for a given round. When the two photons kept by the sender are projected onto the same state (i.e., only a single click is recorded in the four detectors), the two photons sent to Bob have the same polarization. In this case, a dishonest receiver can split the two photons and measure one in each basis. Assuming four detectors with equal efficiencies (which can be guaranteed in practice by appropriate attenuation the higher efficiency ones), and using the fact that for an SPDC source, whenever multiple pairs are produced, there is no correlation among them, we know that the number of undetected multi-photon events is approximately $\frac{1}{3}$ of the number of detected ones. We can then estimate the probability p'_{multi} of an accepted coincidence to be associated with a multi-photon event with Eq. (37).

Note that the statistical check performed by Alice in the second step of the amended protocol (Sec. V B) ensures security under the assumption that there is no coherence in the photon-number basis. This is the case in our implementation, since SPDC produces states of the form $\sum_{n=0}^{\infty} \sqrt{c_n}|n\rangle_1|n\rangle_2$ in the number basis $\{|n\rangle\}$ [54], leaving the individual subsystems in incoherent mixtures of the form $\sum_{n=0}^{\infty} c_n|n\rangle\langle n|$.

To account for the leakage caused by undetected multi-photon emissions to our OT rate expression, we effectively grant Bob an amount of information about Alice's measurement outcomes equal to the number of indices in $I_{\bar{I}_t}$ associated to multi-photon events, upper bounded by $p_{\mathrm{multi}}(1-\alpha)N_0$ for large N_0. Subtracting this leak to the total entropy expression in Eq. (32) leads to a version of Lemma 2 for security against dishonest receiver corrected for the experimental implementation, which differs from the theoretical version by replacing Eq. (13) with

$$\varepsilon'_{\text{exp}} = \sqrt{2}\left(e^{-\frac{1}{2}(1-\alpha)^2 N_{\text{test}}\delta_1^2} + e^{-\frac{1}{2}N_{\text{check}}\delta_1^2}\right)^{\frac{1}{2}} + e^{-D_{\text{KL}}\left(\frac{1}{2}-\delta_2|\frac{1}{2}\right)(1-\alpha)N_0} + \varepsilon_{\text{bind}}(k) + \frac{1}{2}\cdot 2^{\frac{1}{2}\left(n-N_{\text{raw}}\left(\frac{1}{2}-\frac{2\delta_2}{1-2\delta_2}-h\left(\frac{p_{\max}+\delta_1}{\frac{1}{2}-\delta_2}\right)-f\cdot h(p_{\max}+\delta_1)-\frac{p_{\text{multi}}}{\frac{1}{2}-\delta_2}\right)\right)}, \tag{38}$$

VI. DISCUSSION

Using Naor's protocol [35] in conjunction with a linear time OWF (such as a hash function fron the SHA3 or BLAKE family), it is possible to implement the required 2-bit commitment in linear time in k. On the other hand, using an LDPC code with soft-decision decoding and hash-based verification, one can implement an IR scheme which is linear in both the block size N_{raw} (and therefore N_0) and k'. Finally, by taking the universal hash set $\mathbf{F}$ to be the set of Toeplitz matrices of size $N_{\text{raw}} \times n$, and using the FFT algorithm for matrix-vector multiplication, the computation of the output strings can be done in time $O(N_{\text{raw}}\log(N_{\text{raw}}))$. Considering that the protocol requires N_0 commitments and all the remaining computations of random subsets and checks can be done in linear time in N_0, the total protocol running time is $O(N_0(k+k'+\log(N_0))$.

Regarding the practicality of implementing π_{QROT}, the protocol is designed to be compatible with BB84-based QKD setups, both from the physical layer up to the post-processing, only requiring the addition of the commitment scheme. The most important difference to note is that π_{QROT} has significantly lower tolerance for qubit error rate (QBER). While most common QKD protocols can produce keys through QBERs above 10%, this protocol is limited to a maximum of 2.8%. This comparatively reduces the distances at which the protocol can be successful. However, it is important to note that, as opposed to key distribution between trusting parties, there are legitimate use-cases for OT at short range. While being in proximity to each other can help two trusting parties isolate themselves from a third-party eavesdropper, mistrusting parties do not gain anything (security wise) from being in the same place while attempting to do MPC, making the protocol useful regardless of the distance between the users.

Comparisons between classical and quantum protocols can be difficult because physical and technological assumptions, such as access to quantum communication or noisy quantum storage, do not straightforwardly compare with computational hardness assumptions. Furthermore, there is no natural way of quantitatively comparing statistical versus computational security. We can, however, contrast the (dis)advantages of using a computationally secure quantum OT protocol as compared to both fully classical computationally secure protocols, as well as statistically secure quantum ones.

Classical OT protocols based on asymmetric cryptography comprise the overwhelming majority of current real-world implementations of OT. The obvious main advantage of quantum OT is the weaker computational hardness assumption (OWF vs asymmetric cryptography), while the main advantage of current post-quantum classical OT implementations is speed. As shown in Fig. 5, the presented experimental setup is able to produce up to 0.10 OT/s, which pales in comparison to contemporary classical protocols, such as Refs. [30–33], that can achieve upwards of 10^5 OT/s (not including latency between parties) with current off-the-shelf hardware (for more details, see Ref. [33]). This difference can be mitigated by the use of OT extension algorithms, as the difference in speed would only matter during the generation of the base OTs. Note that in this case one should use a OT extension that matches the computational assumption of this work, such as Ref. [55].

Quantum protocols, both discrete variable (DV) [27] and continuous variable (CV) [29], have been shown to achieve statistically secure OT in the quantum noisy-storage model (QNS). Their experimental implementations show comparable values of quantum communication cost in terms of shared signals: 10^8 (no memory encoding assumption), and 10^5 (Gaussian encoding) for CV, and 10^7 for DV. As shown in Fig. 4, our protocol requires 10^6 quantum signals when matching their security ($\varepsilon = 10^{-7}$), which improves upon the alternatives when no additional assumption on the memory encoding of the adversary is made. Less straightforward to compare is the strength of the assumptions of noisy storage and OWFs. We note that the existence of OWFs is an assumption that permeates modern cryptography, from block cipher encryption and message authentication up to public-key cryptography protocols [56], which makes π_{QROT} more suited to be introduced in current cipher suites than protocols with alternative assumptions. In particular, as noted above, OWFs are required for OT extension algorithms. A summary of comparisons between the different approaches can be found in Table II.

Regarding potential improvements and further work, we can identify two main directions to build upon this work: performance and security. Regarding performance, we note that the dominant term in the expression for $\varepsilon_{\max}$ is the one associated with the significance of the parameter estimation [the first term in Eq. (13)]. This translates into the

TABLE II. Comparison of our work with other approaches for OT. N denotes the respective security parameter. Acronyms for assumptions are as follows: OWF, one-way functions; QNS, quantum noisy storage; DDH, decisional Diffie-Helmann; RLWE, ring learning with errors; SLS, space-like separation enforced. Protocols marked with superscript a do not have a reference experimental implementation at the time of writing.

Protocol	Type	Assumption	Quantum Cost	Security
This work	Quantum discrete variable	OWF	$\mathcal{O}(N)$	Indistinguishability UC ROM
GLSV21 [38][a]	Quantum discrete variable	OWF	$Poly(N)$	Stand-alone plain model
S10 [27,48]	Quantum discrete variable	QNS	$\mathcal{O}(N)$	Indistinguishability
FGSPSW18 [29]	Quantum continuous variable	QNS	$\mathcal{O}(N)$	Indistinguishability
MR19 [30]	Classical	DDH	–	Stand-alone ROM
BFGMMS21 [33]	Classical	RLWE	–	UC ROM
P16 [57][a]	Quantum and relativistic discrete variable	SLS	$\mathcal{O}(N)$	Other

relatively large values of N_0 needed to achieve adequate security, which was the bottleneck in the performance of our implementation. One way to reduce the number of signals needed per OT is to modify the protocol to perform many concurrent ROTs in a single run. This would mean performing one single estimation, albeit of a larger sample, that would work for many OTs in such a way that the required number of signals per ROT is decreased. On the topic of increasing security two main directions come to mind. First, we can consider the constructions of *collapsing* hash functions proposed in Refs. [58,59] to implement statistically hiding, computationally collapse binding commitments, which in turn allow for OT protocols that feature forward security (the OT remains secure even if the underlying hash function can be attacked after the commit and open phase of the protocol). The second direction would be a deeper exploration of the composable security of the protocol in the ROM. This can come from generalizing Theorem 2 for any weakly interactive commitments (currently the proof applies only to the LRV25 construction), or applying the techniques developed in Ref. [60] to prove UC security of commitments in the quantum ROM to remove the adversary's limitation of classical access to the oracle. From the practical implementation perspective, it seems natural to integrate quantum OT into both QKD setups for a unified physical layer capable of providing secure communication and computation powered by OT extension and MPC algorithms, bringing the benefits of quantum OT closer to real world usage.

ACKNOWLEDGMENTS

M.L., P.M., and N.P. acknowledge funding from FCT/MECI through national funds and co-funded EU funds under UID/50008: Instituto de Telecomunicações. M.L. acknowledges PhD scholarship PD/BD/114334/2016 as well as FCT project QuantumPrime PTDC/EEI-TEL/8017/2020. N.P. acknowledges the FCT Estímulo ao Emprego Científico Grant No. CEECIND/04594/2017/CP1393/CT000. P.S., M.B. and P.W. acknowledge funding from the European Union's Horizon Europe research and innovation program under Grant Agreement No. 101114043 (QSNP), along with the Austrian Science Fund FWF 42 through [F7113] (BeyondC) and the AFOSR via FA9550-21-1-0355 (QTRUST). D.E. was supported by the JST Moonshot R&D program under Grant JPMJMS226C as well as Japan's Council for Science, Technology and Innovation (CSTI) under the Cross-ministerial Strategic Innovation Promotion Program (SIP) for "Promoting the application of advanced quantum technology platforms to social issues" (Grant JPJ012367). M.G. acknowledges FCT Portugal financing Refs. UIDB/50021/2020 and UIDP/50021/2020 (resp. DOI 10.54499/UIDB/50021/2020 and 10.54499/UIDP/50021/2020).

DATA AVAILABILITY

The data that support the findings of this article are openly available [46].

APPENDIX A: PRELIMINARIES

1. Quantum computational efficiency and distinguishability

We model the quantum capabilities of parties through programs running on quantum computers, for which we adopt a model based on deterministic-control quantum Turing machines [61]. For the purposes of the following definitions, a quantum computer is a device that has a classical interface and a quantum part, which contains the quantum memory registers available to the party. The classical interface has the capabilities of a classical computer augmented with the ability to perform a predefined universal set of quantum operators on the quantum memory registers and perform measurements in the canonical (computational) basis. Given a specified type of quantum computer, a quantum program is a classical description of a set of instructions to be run by the computer, including the quantum operations and measurements to be executed in the quantum part, as well as any classical computation. Quantum programs can be compared with probabilistic classical programs as they both have natural numbers as inputs and outputs. When a quantum computer runs the

program T with input $x \in \mathbb{N}$, we assume that the quantum part of the computer starts with some predefined initial state, performs a sequence of operations on its quantum registers, and upon halting, it outputs $T(x) \in \mathbb{N}$ on its classical interface by reading the appropriate registers associated with the program's output. Each execution of a quantum program is then associated to a quantum operation, which is the result of all the operations performed on the quantum part during the execution of the program.

Definition A1 (Computational efficiency). Let T be a quantum program. We say that T is computationally efficient (or polynomial time) if there exists a polynomial P such that the running time of $T(x)$ is $O(P(x))$.

Definition A2 (Distinguishing advantage). Let X_1, X_2 be two random variables with values in $\mathbb{N}$. For any quantum program T, the *distinguishing advantage* of X_1, X_2 using T is defined as

$$\mathrm{Adv}_T(X_1, X_2) = |\Pr[T(X_1) = 1] - \Pr[T(X_2) = 1]|\,. \quad \text{(A1)}$$

Analogously, let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$. For any quantum program T, the *distinguishing advantage* of ρ_1, ρ_2 using T is defined as

$$\mathrm{Adv}_T(\rho_1, \rho_2) = \left|\Pr[T^{(\rho_1)} = 1] - \Pr[T^{(\rho_2)} = 1]\right|, \quad \text{(A2)}$$

where $T^{(\rho)}$ denotes the classical output of the program starting with the quantum state ρ and zero classical input.

Definition A3 (Indistinguishability-finite). Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ and $\varepsilon \geq 0$. We say that ρ_1 and ρ_2 are ε-indistinguishable, denoted by $\rho_1 \approx_\varepsilon \rho_2$, whenever

$$\mathrm{Adv}_T(\rho_1, \rho_2) \leq \varepsilon, \quad \text{for all quantum programs } T. \quad \text{(A3)}$$

ε-indistinguishability for random variables is defined analogously.

As the following proposition states, to show that two states are ε-indistinguishable, it is enough to upper bound their trace distance D. (For more detail on the relationship of these quantities, see Refs. [62,63].)

Proposition A1. For any pair of quantum states $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ it holds that

$$\rho_1 \approx_{D(\rho_1,\rho_2)} \rho_2. \quad \text{(A4)}$$

Definition A4 (Indistinguishability-asymptotical). Let $\{\rho_1^{(k)} \in \mathcal{D}(\mathcal{H}_k)\}$ and $\{\rho_2^{(k)} \in \mathcal{D}(\mathcal{H}_k)\}$ be two families of density operators. We say that the two families are *statistically indistinguishable* if there exists a negligible function $\varepsilon(k) \geq 0$ such that

$$\rho_1^{(k)} \approx_{\varepsilon(k)} \rho_2^{(k)} \quad \text{for all } k \in \mathbb{N}. \quad \text{(A5)}$$

Furthermore, we say the two families are *computationally indistinguishable* if for every efficient quantum program T, there exists a negligible function $\varepsilon_T(k) \geq 0$ such that

$$\mathrm{Adv}_T\left(\rho_1^{(k)}, \rho_2^{(k)}\right) \leq \varepsilon_T(k) \quad \text{for all } k \in \mathbb{N}. \quad \text{(A6)}$$

Statistical and computational indistinguishability for random variables is defined analogously.

Recall from Sec. II that, when the parameter k is implicit, we may omit the explicit dependence on k and use $\approx$ and $\approx^{(c)}$ for statistical and computational indistinguishability, respectively. We now turn our attention to the properties of indistinguishable states. It is worth noting that computational indistinguishability is only meaningful in terms of information security when the adversary is assumed to have limited computational capabilities. It is important then to define the type of quantum operations such an adversary can perform.

Definition A5 (Efficient quantum operation). We say that a family $\{\mathcal{E}^{(k)}\}_{k=1}^{\infty}$ of quantum operations is efficient if there exists an efficient quantum program T such that, for each k, $\mathcal{E}^{(k)}$ is the associated operation applied to the quantum part of the machine while running T on input k.

The following properties are straightforward to prove from Definitions A3 and A4 and the properties of trace distance:

Lemma A1 (Properties of indistinguishable states I). Let $\rho_1, \rho_2, \rho_3 \in \mathcal{D}(\mathcal{H})$:

(1) $\rho_1 \approx_\varepsilon \rho_2 \wedge \rho_2 \approx_{\varepsilon'} \rho_3 \Rightarrow \rho_1 \approx_{\varepsilon+\varepsilon'} \rho_3$.
(2) $\rho_1 \approx_\varepsilon \rho_2 \wedge \sigma_1 \approx_{\varepsilon'} \sigma_2 \Rightarrow \rho_1 \otimes \sigma_1 \approx_{\varepsilon+\varepsilon'} \rho_2 \otimes \sigma_2$.
(3) Let $x \in \mathcal{X}$. For any probability distribution P_x, assume that $(\forall x \in \mathcal{X})\ \rho_1^x \approx_{\varepsilon^x} \rho_2^x$. Then

$$\sum_{x\in\mathcal{X}} P_x \rho_1^x \approx_{\varepsilon^{\max}} \sum_{x\in\mathcal{X}} P_x \rho_2^x \quad \text{where } \varepsilon^{\max} = \max_{x\in\mathcal{X}}\{\varepsilon^x\}.$$

(4) $\rho_1 \approx_\varepsilon \rho_2 \Rightarrow \mathcal{E}(\rho_1) \approx_\varepsilon \mathcal{E}(\rho_2)$, for any completely positive, trace non-increasing map $\mathcal{E}$.

Lemma A2 (Properties of indistinguishable states II). Let $\{\rho_1(k)\}, \{\rho_2(k)\}, \{\rho_3(k)\}$ be families of density operators parameterized by $k = 1, 2, \ldots$ The following statements hold for asymptotic computational indistinguishability:

(1) $\rho_1 \approx^{(c)} \rho_2 \wedge \rho_2 \approx^{(c)} \rho_3 \Rightarrow \rho_1 \approx^{(c)} \rho_3$.
(2) $\rho_1 \approx^{(c)} \rho_2 \wedge \sigma_1 \approx^{(c)} \sigma_2 \Rightarrow \rho_1 \otimes \sigma_1 \approx^{(c)} \rho_2 \otimes \sigma_2$.
(3) Let $x \in \mathcal{X}$. For any probability distribution P_x, assume that $(\forall x \in \mathcal{X})\ \rho_1^x \approx^{(c)} \rho_2^x$. Then

$$\sum_{x\in\mathcal{X}} P_x \rho_1^x \approx^{(c)} \sum_{x\in\mathcal{X}} P_x \rho_2^x.$$

(4) $\rho_1 \approx^{(c)} \rho_2 \Rightarrow \mathcal{E}(\rho_1) \approx^{(c)} \mathcal{E}(\rho_2)$, where $\{\mathcal{E}^{(k)}\}$ is an efficient family of quantum operations acting on the respective $\rho_i(k)$.

2. Entropic quantities

We start off by defining a useful pair of quantities for measuring information in quantum systems: the max-entropy and the conditional min-entropy. The max-entropy is a measure of the number of possible different outcomes that can result from measuring a quantum state, whereas the conditional min-entropy is a way of measuring the information that a party can infer from a quantum system given access to another correlated quantum system. This measure will be useful to bound the distance between states based on their internal correlations.

Definition A6 (Max-entropy). Let $\rho \in \mathcal{D}(\mathcal{H})$. The max-entropy of ρ is defined as

$$H_{\max}(\rho) = \log\left(\dim(\mathrm{supp}(\rho))\right), \tag{A7}$$

where $\mathrm{supp}(\rho)$ denotes the support subspace of ρ and dim denotes its dimension.

Definition A7 (Min-entropy and conditional min-entropy). Let $\rho \in \mathcal{D}(\mathcal{H})$ and $\lambda_{\max}(\rho)$ denote the maximum eigenvalue of ρ. The min-entropy of ρ is defined as

$$H_{\min}(\rho) = -\log(\lambda_{\max}(\rho)). \tag{A8}$$

Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\sigma_B \in \mathcal{D}(\mathcal{H}_B)$. The conditional min-entropy of ρ_{AB} given σ_B is defined as

$$H_{\min}(\rho_{AB} \mid \sigma_B) = -\log(\lambda_{\sigma_B}), \tag{A9}$$

where λ_{σ_B} is the minimum real number such that $\lambda_{\sigma_B}(\mathbb{1}_A \otimes \sigma_B) - \rho_{AB}$ is non-negative. The conditional min-entropy of ρ_{AB} given $\mathcal{H}_B$ is defined as

$$H_{\min}(A \mid B)_\rho = \sup_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} H_{\min}(\rho_{AB} \mid \sigma_B). \tag{A10}$$

Furthermore, let $\varepsilon > 0$. The ε-smooth conditional min-entropy is defined as

$$H^{\varepsilon}_{\min}(A \mid B)_\rho = \sup_{\rho'_{AB} \in \mathcal{B}^{\varepsilon}(\rho_{AB})} H_{\min}(A \mid B)_{\rho'}, \tag{A11}$$

where $\mathcal{B}^{\varepsilon}(\rho_{AB}) = \{\rho'_{AB} : D(\rho_{AB}, \rho'_{AB}) < \varepsilon\}$.

The smooth conditional min-entropy is in general hard to compute. Because of this, it is useful to have some tools to bound it for states that have some specific forms. In our case we are interested in states that are partially *classical*.

Definition A8 (Partially classical states). A quantum state described by the density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is classical in $\mathcal{H}_A$ (or classical in A) if it can be written in the form

$$\rho_{AB} = \sum_x \lambda_x |x\rangle\langle x|_A \otimes \rho^x_B, \tag{A12}$$

where the set $\{|x\rangle\}_x$ is an orthonormal basis for $\mathcal{H}_A$. A multipartite state is said to be classical if it is classical in all its parts.

When dealing with partially classical states as shown in Eq. (A12), we will refer to the operators ρ^x_B as the state of the system B *conditioned* to x.

Lemma A3 (Properties of min- and max-entropy). Let $\varepsilon, \varepsilon' \geq 0$:

(1) $H_{\min}(\rho_A \otimes \rho_B|\rho_B) = -\log(\lambda_{\max}(\rho_A))$.
(2) $H^{\varepsilon+\varepsilon'}_{\min}(AA' \mid BB')_{\rho\otimes\rho'} \geq H^{\varepsilon}_{\min}(A \mid B)_\rho + H^{\varepsilon'}_{\min}(A' \mid B')_{\rho'}$.
(3) $H^{\varepsilon}_{\min}(A \mid BC)_\rho \leq H^{\varepsilon}_{\min}(A \mid B)_\rho$.
(4) $H^{\varepsilon}_{\min}(AB \mid C)_\rho \leq H^{\varepsilon}_{\min}(A \mid BC)_\rho + H_{\max}(\rho_B)$.
(5) $H^{\varepsilon}_{\min}(A \mid B)_\rho \geq \inf_x\{H^{\varepsilon}_{\min}(\rho^x_A)\}$, whenever the state ρ_{AB} is classical on B.

We use universal hashing to implement randomness extraction in the final steps of the protocol. The proof both Lemmas A3 and A4 can be found in Ref. [64].

Definition A9 (Universal hashing). A set of functions $\mathbf{F} = \{f_i : \{0,1\}^m \to \{0,1\}^n\}$ is a universal hash family if, for all $x, y \in \{0,1\}^m$, such that $x \neq y$, and i chosen uniformly at random, we have

$$\Pr[f_i(x) = f_i(y)] \leq \frac{1}{2^n}. \tag{A13}$$

Lemma A4 (Quantum leftover hash). Let $\mathbf{F} = \{f_i : \{0,1\}^n \to \{0,1\}^\ell\}$ be a universal hash family, let $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_F, \mathcal{H}_E$ be Hilbert spaces such that $\{|x\rangle\}_{x\in\{0,1\}^n}, \{|f_i\rangle\}_{f_i\in\mathbf{F}}$, and $\{|e\rangle\}_{e\in\{0,1\}^\ell}$ are orthonormal bases for $\mathcal{H}_A, \mathcal{H}_F$, and $\mathcal{H}_E$ respectively. Then for any $\varepsilon \geq 0$ and any state of the form

$$\rho_{ABFE} = \frac{1}{|\mathbf{F}|} \sum_{\substack{x\in\{0,1\}^n \\ f_i\in\mathbf{F}}} \left(\lambda(x)|f_i\rangle\langle f_i|_F |f_i(x)\rangle\langle f_i(x)|_E \otimes |x\rangle\langle x|_A \rho^x_B\right), \tag{A14}$$

it holds that

$$\rho_{EBF} \approx_{\varepsilon'} \mathsf{U}_E \otimes \rho_{BF}, \tag{A15}$$

with

$$\varepsilon' = \varepsilon + \frac{1}{2} \cdot 2^{-\frac{1}{2}(H^{\varepsilon}_{\min}(A|B)_\rho - \ell)}. \tag{A16}$$

APPENDIX B: DETAILED PROOF OF THEOREM 1

1. Supporting lemmas

One of the main features to analyze for the security against a dishonest receiver is the potential information that he can learn about the sender's strings given the quantum state that remains with him after the commit and open phase. In order to talk about the security of the protocol independently of the specific cheating strategy that may be used by the dishonest parties (or possible effects that the environment can have in the shared quantum state), we want to understand the properties that a quantum state that passes Alice's test at step (7) can have. We do this through the following lemma, a version of which was originally proven in Ref. [40]. Here, we provide a more self-contained statement and make explicit the trace distance bound.

Lemma B1. Let $\varepsilon > 0$, $I = \{1, \dots, N\}$, and $\rho_{T,\hat{X},X,E}$ be a density operator of the form

$$\begin{aligned}\rho_{T,\hat{X},X,E} &= \sum_{\substack{I_1,I_2 \\ \in\mathcal{P}(I)\setminus\{\emptyset\}}} q(I_1,I_2)|I_1,I_2\rangle\langle I_1,I_2|_T \\ &\quad \otimes |\hat{x}\rangle\langle\hat{x}|_{\hat{X}} \otimes |\psi\rangle\langle\psi|_{X,E} \\ |\psi\rangle_{X,E} &= \sum_{x\in\{0,1\}^N} \beta^x |x\rangle_X |\phi^x\rangle_E,\end{aligned} \tag{B1}$$

where $\dim(\mathcal{H}_{\hat{X}}) = \dim(\mathcal{H}_X) = 2^N$ and $\mathcal{P}(I)$ denotes the set of all subsets of I. For each I_1, I_2 define the set

$$B_{I_1,I_2} = \{x \in \{0,1\}^N : |r_H(x_{I_1} \oplus \hat{x}_{I_1}) - r_H(x_{I_2} \oplus \hat{x}_{I_2})| < \varepsilon\}. \tag{B2}$$

Additionally, let $\mathcal{Q}(N,\varepsilon)$ be a function such that, whenever the subsets I_1, I_2 are sampled according to q, it holds that

$$\Pr[|r_H(x_{I_1} \oplus \hat{x}_{I_1}) - r_H((x_{I_2} \oplus \hat{x}_{I_2})| > \varepsilon] \le \mathcal{Q}(N,\varepsilon) \tag{B3}$$

independently of x. There exists a state $\tilde{\rho}_{T,\hat{X},X,E}$ of the form

$$\begin{aligned}\tilde{\rho}_{T,\hat{X},X,E} &= \sum_{\substack{I_1,I_2 \\ \in\mathcal{P}(I)\setminus\{\emptyset\}}} q(I_1,I_2)|I_1,I_2\rangle\langle I_1,I_2|_T \\ &\quad \otimes |\hat{x}\rangle\langle\hat{x}|_{\hat{X}} \otimes |\psi_{I_1,I_2}\rangle\langle\psi_{I_1,I_2}|_{X,E} \\ |\psi_{I_1,I_2}\rangle &= \sum_{x\in B_{I_1,I_2}} \tilde{\beta}^x_{I_1,I_2} |x\rangle_X |\phi^x_{I_1,I_2}\rangle_E,\end{aligned} \tag{B4}$$

such that

$$D(\rho_{T,\hat{X},X,E}, \tilde{\rho}_{T,\hat{X},X,E}) \le \mathcal{Q}(N,\varepsilon)^{\frac{1}{2}}. \tag{B5}$$

Proof. First, we choose an adequate definition for the $\tilde{\beta}^x_{I_1,I_2}$ and then show that, under that choice, the bound in Eq. (B5) holds. Note that we can write the state

$$\begin{aligned}|\psi\rangle_{X,E} &= \sum_{x\in\{0,1\}^N} \beta_x |x\rangle_X |\phi^x\rangle_E \\ &= \underbrace{\left(\sum_{x\in B_{I_1,I_2}} |\beta_x|^2\right)^{\frac{1}{2}}}_{\lambda_{I_1,I_2}} \underbrace{\frac{\sum_{x\in B_{I_1,I_2}} \beta_x |x\rangle_X |\phi^x\rangle_E}{\left(\sum_{x\in B_{I_1,I_2}} |\beta_x|^2\right)^{\frac{1}{2}}}}_{|\psi_{I_1,I_2}\rangle_{X,E}} + \underbrace{\left(\sum_{x\notin B_{I_1,I_2}} |\beta_x|^2\right)^{\frac{1}{2}}}_{\lambda^{\perp}_{I_1,I_2}} \underbrace{\frac{\sum_{x\notin B_{I_1,I_2}} \beta_x |x\rangle_X |\phi^x\rangle_E}{\left(\sum_{x\notin B_{I_1,I_2}} |\beta_x|^2\right)^{\frac{1}{2}}}}_{|\psi^{\perp}_{I_1,I_2}\rangle_{X,E}}.\end{aligned} \tag{B6}$$

The trace distance between the pure states $|\psi\rangle$ and $|\psi_{I_1,I_2}\rangle$ is given by $\sqrt{1 - |\langle\psi|\psi_{I_1,I_2}\rangle|^2} = \lambda^{\perp}_{I_1,I_2}$, hence the trace distance between the complete joint states is given by

$$\begin{aligned}D(\rho_{T,\hat{X},X,E}, \tilde{\rho}_{T,\hat{X},X,E})^2 &\le \left(\sum_{I_1,I_2\in\mathcal{T}(N,\alpha)} q D(|\psi\rangle\langle\psi|_{X,E}, |\psi_{I_1,I_2}\rangle\langle\psi_{I_1,I_2}|_{X,E})\right)^2 \\ &= \left(\sum_{I_1,I_2\in\mathcal{T}(N,\alpha)} q\lambda^{\perp}_{I_1,I_2}\right)^2 \\ &\le \sum_{I_1,I_2\in\mathcal{T}(N,\alpha)} q{\lambda^{\perp}_{I_1,I_2}}^2,\end{aligned} \tag{B7}$$

where Jensen's inequality was used in the last step. We proceed now to bound the right side of Eq. (B7). For that purpose consider the function

$$\xi(I_1, I_2, x) = \begin{cases} 0, & \text{if} \quad x \in B_{I_1,I_2} \\ 1, & \text{otherwise} \end{cases}, \tag{B8}$$

so that

$$\begin{aligned} &\sum_{I_1,I_2} q(I_1,I_2)\xi(I_1,I_2,x) \\ &\quad = \Pr[|r_H(\hat{x} \oplus x|_{I_1}) - r_H(\hat{x} \oplus x|_{I_2})| > \varepsilon] \\ &\quad = \mathcal{Q}(N,\varepsilon). \end{aligned} \tag{B9}$$

Hence,

$$\begin{aligned} &\|\rho_{T,\hat{X},X,E} - \tilde{\rho}_{T,\hat{X},X,E}\|^2 \\ &\quad \le \sum_{I_1,I_2} q(I_1,I_2) {\lambda^{\perp}_{I_1,I_2}}^2 \\ &\quad = \sum_{I_1,I_2} q(I_1,I_2) \sum_{x \notin B_{I_1,I_2}} |\beta_x|^2 \\ &\quad = \sum_{I_1,I_2} q(I_1,I_2) \sum_{x \in \{0,1\}^N} \xi(I_1,I_2,x)|\beta_x|^2 \\ &\quad = \sum_{x \in \{0,1\}^N} |\beta_x|^2 \sum_{I_1,I_2} q(I_1,I_2)\xi(I,x) \\ &\quad \le \sum_{x \in \{0,1\}^N} |\beta_x|^2 \mathcal{Q}(N,\varepsilon) = \mathcal{Q}(N,\varepsilon), \end{aligned} \tag{B10}$$

as required. ■

In order to use the above result in the context of the π_{QROT} protocol, we need to find an appropriate function $\mathcal{Q}(N,\varepsilon)$ that satisfies Eq. (B3) for the case when the $\hat{x}, x$ are the respective measurement outcomes of Alice and Bob when measuring in the same basis. We do this through the following lemma based on the Hoeffding inequality for sampling without replacement.

Definition B1. Given a set I and an integer $n \le |I|$, define the set $\mathcal{T}(n,I)$ as the set of all subsets of I with size n.

Lemma B2 (Hoeffding's inequalities). Let $x \in \{0,1\}^N$, $\delta > 0, I = \{1,\dots,N\}$ and $0 < \alpha < \frac{1}{2}$ such that $\alpha N \in \mathbb{N}$.

(a) (Inequality for sampling without replacement comparing the sampled subset with the whole set) For $I_t \in \mathcal{T}(\alpha N, I)$ sampled uniformly, it holds that

$$\Pr[|r_H(x|_{I_t}) - r_H(x)| > \delta] \le 2e^{-2\alpha N\delta^2}. \tag{B11}$$

(b) (Inequality for sampling without replacement comparing the sampled subset with its complement)

$$\Pr[|r_H(x|_{I_t}) - r_H(x|_{\bar{I}_t})| > \delta] \le 2e^{-2\alpha(1-\alpha)^2 N\delta^2}. \tag{B12}$$

(c) (Inequality for sampling without replacement comparing the sampled subset with its complement and ignoring part of the sample) Let $n \in \{n_0,\dots,\alpha N\}$ be sampled according to some distribution $P(n)$. For $I_s \in \mathcal{T}(n, I_t)$ sampled uniformly, it holds that

$$\begin{aligned} &\Pr[|r_H(x|_{I_s}) - r_H(x|_{\bar{I}_t})| > \delta] \\ &\quad \le 2\left(e^{-\frac{1}{2}\alpha(1-\alpha)^2 N\delta^2} + e^{-\frac{1}{2}n_0\delta^2}\right). \end{aligned} \tag{B13}$$

Proof.

(a) This is the original Hoeffding inequality for sampling without replacement, the proof of which can be found in Ref. [65].

(b) Note that we can write $r_H(x) = \alpha r_H(x|_{I_t}) + (1-\alpha) r_H(x|_{\bar{I}_t})$. Substituting $r_H(x)$ in Eq. (B11) we get

$$\begin{aligned} &\Pr[|r_H(x|_{I_t}) - \alpha r_H(x|_{I_t}) + (1-\alpha) r_H(x|_{\bar{I}_t})| > \delta'] \\ &\quad = \Pr[|r_H(x|_{I_t}) - r_H(x|_{\bar{I}_t})| > \delta'/(1-\alpha)] \end{aligned} \tag{B14}$$

$$\le 2e^{-2\alpha N\delta'^2}. \tag{B15}$$

The result is obtained by taking $\delta' = (1-\alpha)\delta$

(c) Let us consider first the case where n is fixed. From the triangle inequality we know that

$$|r_H(x|_{I_{s(n)}}) - r_H(x|_{\bar{I}_t})| > \delta \Rightarrow |r_H(x|_{I_{s(n)}}) - r_H(x|_{I_t})| + |r_H(x|_{I_t}) - r_H(x|_{\bar{I}_t})| > \delta \tag{B16}$$

$$\Rightarrow |r_H(x|_{I_{s(n)}}) - r_H(x|_{I_t})| > \delta/2 \vee |r_H(x|_{I_t}) - r_H(x|_{\bar{I}_t})| > \delta/2, \tag{B17}$$

and hence, by the union bound

$$\Pr[|r_H(x|_{I_{s(n)}}) - r_H(x|_{\bar{I}_t})| > \delta] \le \Pr[|r_H(x|_{I_{s(n)}}) - r_H(x|_{I_t})| + |r_H(x|_{I_t}) - r_H(x|_{\bar{I}_t})| > \delta] \tag{B18}$$

$$\le \Pr[|r_H(x|_{I_{s(n)}}) - r_H(x|_{I_t})| > \delta/2] + \Pr[|r_H(x|_{I_t}) - r_H(x|_{\bar{I}_t})| > \delta/2] \tag{B19}$$

$$\le 2e^{-\frac{1}{2}n\delta^2} + 2e^{-\frac{1}{2}\alpha(1-\alpha)^2 N\delta^2}, \tag{B20}$$

where the last expression comes from applying the (b) and (a) inequalities to the first and second terms of Eq. (B19) respectively. Using this, we can consider the case in which n is not fixed, but instead follows a probability distribution $P(n)$ such that $P(n) = 0$ for $n < n_0$. For this case

$$\Pr[|r_H(x|_{I_s}) - r_H(x|_{\bar{I}_t})| > \delta] = \sum_n P(n) \Pr[|r_H(x|_{I_{s(n)}}) - r_H(x|_{\bar{I}_t})| > \delta] \tag{B21}$$

$$\leq \sum_n P(n) 2 \left(e^{-\frac{1}{2}\alpha(1-\alpha)^2 N\delta^2} + e^{-\frac{1}{2}n\delta^2} \right) \tag{B22}$$

$$\leq 2 \sum_n P(n) \left(e^{-\frac{1}{2}\alpha(1-\alpha)^2 N\delta^2} + e^{-\frac{1}{2}n_0\delta^2} \right) \tag{B23}$$

$$= 2 \left(e^{-\frac{1}{2}\alpha(1-\alpha)^2 N\delta^2} + e^{-\frac{1}{2}n_0\delta^2} \right). \tag{B24}$$

■

The following lemma helps us bound the conditional min-entropy of a partially measured pure state by comparing it with the one of an appropriately chosen, partially measured mixed state. A proof of this result can be found in Ref. [66].

Lemma B3 (Entropy bound for post-measurement states). Let $\mathcal{H}_A$ and $\mathcal{H}_E$ be Hilbert spaces and $\{|x\rangle\}_{x\in\mathcal{X}}$, $\{|y\rangle\}_{y\in\mathcal{Y}}$ be orthonormal bases for $\mathcal{H}_A$. Let $J \subseteq \mathcal{X}$, define the states

$$\rho_{AE} = |\phi\rangle\langle\phi|_{AE} \quad \text{with } |\phi\rangle_{AE} = \sum_{x\in J} \beta_x |x\rangle_A |\phi^x\rangle_E, \tag{B25}$$

$$\rho^{\text{mix}}_{AE} = \sum_{x\in J} |\beta_x|^2 |x\rangle\langle x|_A \otimes |\phi^x\rangle\langle\phi^x|_E. \tag{B26}$$

Denote by σ_{YE} and σ^{mix}_{YE} the states resulting from measuring the subsystem A of ρ_{AE} and ρ^{mix}_{AE} respectively in the basis $\{|y\rangle\}_{y\in\mathcal{Y}}$, storing the result in the system Y, and then tracing out the A subsystem; then it holds that

$$H_{\min}(Y \mid E)_\sigma \geq H_{\min}(Y \mid E)_{\sigma^{\text{mix}}} - \log |J|. \tag{B27}$$

2. Proof of Lemma 3

Here we present a proof of Lemma 3 used in the protocol's correctness analysis in Sec. IV.

Lemma B4. Let $X^A_{I_0}, X^A_{I_1}, C, Y^B$ denote the systems holding the information of the respective values $x^A_{I_0}, x^A_{I_1}, c$, and y^B of π_{QROT}. Denote by $\rho^\top$ the parties' joint state at the end of step (11) conditioned that Bob constructed the sets (I_0, I_1) during step (9) and the protocol has not aborted. Assume both parties follow the steps of the protocol, then

$$\rho^\top_{X^A_{I_0}, X^A_{I_1}, C, Y^B} \approx_{\varepsilon_{\text{IR}}(k')} \tilde{\rho}^\top_{X^A_{I_0}, X^A_{I_1}, C, Y^B}, \tag{B28}$$

where $\varepsilon_{\text{IR}}(k')$ is a negligible function given by the security of the underlying information reconciliation scheme, k' its associated security parameter, and

$$\tilde{\rho}^\top_{X^A_{I_0}, X^A_{I_1}, C, Y^B} = \frac{1}{2^{(2N_{\text{raw}}+1)}} \sum_{\substack{x_{I_0}, x_{I_1} \\ c}} |x_{I_0}\rangle\langle x_{I_0}|_{X^A_{I_0}} |x_{I_1}\rangle\langle x_{I_1}|_{X^A_{I_1}} |x_{I_0}\rangle\langle x_{I_0}|_{Y^B} |c\rangle\langle c|_C. \tag{B29}$$

Proof. Note that, because the state shared by Alice at step (1) of the protocol is a tensor product of maximally entangled states, the state of Alice's part is a product of maximally mixed states. This means that, regardless of the measurement bases θ^A, the outcome of her measurements x^A is always uniform in $\{0,1\}^{N_0}$. Let $\rho^{(2)}_{\Theta^A, \Theta^B, X^A, X^B}$ be the state of the parties' respective measurement bases and outcomes at the end of step (2) of the protocol, we can write

$$\rho^{(2)}_{\Theta^A, \Theta^B, X^A, X^B} = \frac{1}{2^{2N_0}} \sum_{\theta^A, \theta^B} |\theta^A, \theta^B\rangle\langle\theta^A, \theta^B|_{\Theta^A\Theta^B} \frac{1}{2^{N_0}} \sum_{x^A} |x^A\rangle\langle x^A|_{X^A} \otimes \sum_{x^B} P(x^B \mid x^A, \theta^A, \theta^B) |x^B\rangle\langle x^B|_{X^B}, \tag{B30}$$

where $P(x^B \mid x^A, \theta^A, \theta^B)$ denotes the probabilities of Bob's outcomes given each parties measurement bases and Alice's measurement outcomes, which in turn depends on the effect of the transmission channel when the state was shared from Alice's laboratory. All operations will be classical from this point onwards. To arrive to Eq. (B29), we first need to show that the abort operations within the protocol do not bias the distribution of possible values of x_{I_0} and x_{I_1}, and then use the verifiability property of the IR scheme to ensure that $y^B = x_{I_0}$ with high probability.

We will consider now the two abort instructions at steps (7) and (9) as a single quantum operation $\mathcal{E}$ that maps the state to the zero operator if any of the two abort conditions is satisfied and applies the identity map otherwise. Let us first separate the values of $\theta^{A/B}$ and x^B that "survive" the abort operation. For any given values of x^A and I_t, define the sets

$$J^{(1)}_{I_t,x^A} = \{(\theta^A,\theta^B) : w_H\left(\overline{\theta^A_{I_t} \oplus \theta^B_{I_t}}\right) \geq N_{\text{check}} \wedge N_{\text{raw}} \leq w_H(\theta^A_{I_t} \oplus \theta^B_{I_t}) \leq (1-\alpha)N_0 - N_{\text{raw}}\} \tag{B31}$$

$$J^{(2)}_{I_t,x^A,\theta^A,\theta^B} = \{x^B : r_H(x^A_{I_s} \oplus x^B_{I_s}) \leq p_{\max}\}, \tag{B32}$$

where $w_H(\cdot)$ denotes the Hamming weight function. Let $\mathcal{T} = \mathcal{T}(\alpha N_0, I)$ be the set of all subsets of $I = \{1,\dots,N_0\}$ of size αN_0, and denote by S the system where Bob holds the information of the sets I_0 and I_1. The joint state of the systems $C, X^{A/B}, \Theta^{A/B}, S$ at the end of step (10) of the protocol is

$$\begin{aligned}\rho^{(10)}_{C,X^A,X^B,\Theta^A,\Theta^B,S} = \frac{1}{2}\sum_c |c\rangle\langle c|_C \frac{1}{2^{N_0}}\sum_{x^A}|x^A\rangle\langle x^A|_{X^A} \frac{1}{|\mathcal{T}|\cdot 2^{2N_0}}\sum_{I_t}\sum_{\substack{(\theta^A,\theta^B)\\ \in J^{(1)}}} |\theta^A,\theta^B\rangle\langle\theta^A,\theta^B|_{\theta^A,\theta^B} \\ \otimes \sum_{x^B\in J^{(2)}} P(x^B \mid x^A,\theta^A,\theta^B)|x^B\rangle\langle x^B|_{X^B} \sum_{I_0,I_1} P(I_0,I_1 \mid I_t,\theta^A,\theta^B)|I_0,I_1\rangle\langle I_0,I_1|_S,\end{aligned} \tag{B33}$$

where the conditional distribution $P(I_0,I_1 \mid I_t,\theta^A,\theta^B)$ notably does not depend on x^A or c. Tracing out the $\Theta^{A/B}$ systems and rearranging terms we get

$$\begin{aligned}\rho^{(10)}_{S,C,X^A,X^B} = \sum_{I_0,I_1} P(I_0,I_1)|I_0,I_1\rangle\langle I_0,I_1|_S \\ \otimes \underbrace{\frac{1}{2^{N_0+1}}\sum_{x^A,c}|x^A,c\rangle\langle x^A,c|_{X^A,C}\sum_{x^B} P(x^B \mid I_0,I_1,x^A)|x^B\rangle\langle x^B|_{X^B}}_{\text{conditioned state}\quad \rho^{(10)}_{C,X^A,X^B}(I_0,I_1)},\end{aligned} \tag{B34}$$

where

$$P(I_0,I_1) = \frac{1}{|\mathcal{T}|\cdot 2^{2N_0}}\sum_{I_t}\sum_{\substack{(\theta^A,\theta^B)\\ \in J^{(1)}}} P(I_0,I_1 \mid I_t,\theta^A,\theta^B), \tag{B35}$$

and

$$P(x^B \mid I_0,I_1,x^A) = \frac{\sum_{I_t}\sum_{\substack{(\theta^A,\theta^B)\\ \in J^{(1)}}} P(x^B \mid x^A,\theta^A,\theta^B)P(I_0,I_1 \mid I_t,\theta^A,\theta^B)}{\sum_{I_t}\sum_{\substack{(\theta^A,\theta^B)\\ \in J^{(1)}}} P(I_0,I_1 \mid I_t,\theta^A,\theta^B)}. \tag{B36}$$

Now that we have a form for the conditioned state as pointed out in Eq. (B34), we can move to the action of step (11), where Bob computes $y^B = \mathtt{dec}(\mathtt{syn}(x^A_{I_0}, x^B_{I_0}))$. The resulting state of the systems $C, X^A_{I_0}, X^A_{I_1}, Y^B$ is then given by

$$\rho^{(11)}_{C,X^A_{I_0},X^A_{I_1},Y^B}(I_0,I_1) = \frac{1}{2^{2N_{\text{raw}}+1}}\sum_{\substack{x^A_{I_0},x^A_{I_1}\\ c}} |x^A_{I_0},x^A_{I_1}\rangle\langle x^A_{I_0},x^A_{I_1}|_{X^A_{I_0},X^A_{I_1}}|c\rangle\langle c|_C \otimes \left(P_{\text{correct}}|x^A_{I_0}\rangle\langle x^A_{I_0}|_{Y^B} + P_\perp|\perp\rangle\langle\perp|_{Y^B} + P_{\text{error}}\sigma_{Y^B}\right), \tag{B37}$$

for some coefficients $P_{\text{correct}}, P_\perp, P_{\text{error}}$, and state σ orthogonal to both $|x^A_{I_0}\rangle\langle x^A_{I_0}|$ and $|\perp\rangle\langle\perp|$. By applying the verifiability property of the IR scheme with security parameter k' (where $P_{\text{error}} = \varepsilon_{\text{IR}}(k')$), and conditioning the resulting state to not having aborted, we get the desired result

$$\rho^{(11),\top}_{X^A_{I_0},X^A_{I_1},Y^B,C}(I_0,I_1) \approx_{\varepsilon_{\text{IR}}(k')} \frac{1}{2^{2N_{\text{raw}}+1}}\sum_{\substack{x^A_{I_0},x^A_{I_1}\\ c}} |x^A_{I_0},x^A_{I_1}\rangle\langle x^A_{I_0},x^A_{I_1}|_{X^A_{I_0},X^A_{I_1}}|x^A_{I_0}\rangle\langle x^A_{I_0}|_{Y^B}|c\rangle\langle c|_C. \tag{B38}$$

■

3. Proof of Lemma 4

Here we present a proof of Lemma 4 introduced in Sec. IV B regarding the hiding property of the commitment in the context of π_{QROT}.

Lemma B5. Assuming Bob follows the protocol, for any $J \subseteq I$, the state of the system $A, \mathrm{COM}, \mathrm{OPEN}_J, \Theta_{\bar{J}}^B$ after step (4) satisfies

$$\rho_{A,\mathrm{COM},\mathrm{OPEN}_J,\Theta_{\bar{J}}^B} \approx^{(c)} \rho_{A,\mathrm{COM},\mathrm{OPEN}_J} \otimes \mathsf{U}_{\Theta_{\bar{J}}^B}, \tag{B39}$$

where

$$\mathsf{U}_{\Theta_{\bar{J}}^B} = \frac{1}{2^{N_0-|J|}} \sum_{\theta_{\bar{J}}^B} |\theta_{\bar{J}}^B\rangle\langle\theta_{\bar{J}}^B|_{\Theta_{\bar{J}}^B}, \tag{B40}$$

denotes the uniform distribution over all possible values of $\theta_{\bar{J}}^B$.

Proof. We start by describing the general form of the state prepared by Alice at the beginning of the protocol, which is sent to Bob. Because the value of $r \in \{0,1\}^{n_r}$ sent by Alice in step (3) as part of the commitment scheme is independent of any of Bob's actions, we can consider without loss of generality that it is prepared at the start of the protocol. The state shared at the beginning of the protocol (after Bob receives his qubit shares) has a general form given by

$$\begin{aligned}\rho_{\Phi^B,R,A}^{(0)} &= |\psi^{(0)}\rangle\langle\psi^{(0)}|_{\Phi^B,R,A} \quad |\psi^{(0)}\rangle_{\Phi^B,R,A} \\ &= \sum_{x,r} \alpha^{x,r} |x\rangle_{\Phi^B} |r\rangle_R |\phi^{x,r}\rangle_A,\end{aligned} \tag{B41}$$

where the $|\phi^{x,r}\rangle$ are not necessarily orthogonal. Using the Hadamard operator H, we can define the states

$$|x,\theta\rangle = H^\theta |x\rangle = H^{\otimes\theta_1} \otimes \ldots \otimes H^{\theta_{N_0}} |x\rangle, \tag{B42}$$

and write, for any string of basis choices $\theta \in \{0,1\}^{N_0}$, the state $|\psi^{(0)}\rangle_{\Phi^B,R,A}$ as

$$\begin{aligned}|\psi^{(0)}\rangle_{\Phi^B,R,A} &= \sum_{x',r} \alpha^{x',r} |\phi^{x',r}\rangle_A |r\rangle_R \sum_x \langle x|H^\theta|x'\rangle |x,\theta\rangle_{\Phi^B} \\ &= \sum_{x,r} \beta^{x,\theta,r} |x,\theta\rangle_{\Phi^B} |r\rangle_R |\phi^{x,\theta,r}\rangle_A,\end{aligned} \tag{B43}$$

with

$$\begin{aligned}\beta^{x,\theta,r} &= \left(\sum_{x'} |\langle x|H^\theta|x'\rangle \alpha^{x',r}|^2\right)^{\frac{1}{2}} \quad |\phi^{x,\theta,r}\rangle \\ &= \left(\beta^{x,\theta,r}\right)^{-1} \sum_{x'} \langle x|H^\theta|x'\rangle \alpha^{x',r} |\phi^{x',r}\rangle.\end{aligned} \tag{B44}$$

After uniformly sampling the values of θ^B, Bob proceeds to perform his measurement on his qubit shares. Let $\mathcal{H}_{X^B}$ denote the system where Bob records the outcome string x^B. Additionally, at step (3) Bob receives the value of r, this is a classical message, which we model as Bob receiving the $\mathcal{H}_R$ system and measuring it in the computational basis upon arrival. We can now easily use Eq. (B43) to get the post-measurement state at the end of step (3) after tracing out $\mathcal{H}_{\Phi^B}$, which is given by

$$\rho_{X^B,\Theta^B,R,A}^{(1)} = \frac{1}{2^{N_0}} \sum_{x^B,\theta^B,r} |\beta^{x^B,\theta^B,r}|^2 |\theta^B\rangle\langle\theta^B|_{\Theta^B} |x^B\rangle\langle x^B|_{X^B} |r\rangle\langle r|_R |\phi^{x^B,\theta^B,r}\rangle\langle\phi^{x^B,\theta^B,r}|_A. \tag{B45}$$

Before proceeding with the protocol, it will be useful to state some basic properties of the above state. Note that even though each of the $|\phi^{x^B,\theta^B,r}\rangle$ depends on θ^B, the partial trace

$$\begin{aligned}&\mathrm{Tr}_{X^B}[\rho^{(1)}] \\ &= \frac{1}{2^{N_0}} \sum_{\theta^B} |\theta^B\rangle\langle\theta^B|_{\Theta^B} \otimes \sum_{x,r} |\alpha^{x,r}|^2 |r\rangle\langle r|_R |\phi^{x,r}\rangle\langle\phi^{x,r}|_A\end{aligned} \tag{B46}$$

has a product form. Furthermore, because honest Bob measures each of his qubits independently, for any $I' \subseteq I$, the partial trace

$$\mathrm{Tr}_{X_{I'}^B}[\rho^{(1)}] = \frac{1}{2^{|I'|}} \sum_{\substack{\theta_{I'}^B \\ \in\{0,1\}^{|I'|}}} |\theta_{I'}^B\rangle\langle\theta_{I'}^B|_{\Theta_{I'}^B} \otimes \rho_{X_{I'}^B,\Theta_{I'}^B,R,A}^{(1)} \tag{B47}$$

also has a product form. As Alice will be able to perform quantum operations on her part of the joint state, it is important to note that the above property holds even after the A subsystem undergoes an arbitrary CPTP transformation independent of $\Theta_{I'}^B$ and $X_{I'}^B$. During step (4) Bob commits his values of θ^B and x^B, for that he samples the values of $s = (s_1, \ldots, s_{N_0})$ and computes

$$\begin{aligned} \mathrm{com} &= \left(\mathtt{com}((\theta_1^B, x_1^B), s_1, r), \ldots, \mathtt{com}((\theta_{N_0}^B, x_{N_0}^B), s_{N_0}, r)\right), \\ \mathrm{open} &= \left(\mathtt{open}((\theta_1^B, x_1^B), s_1), \ldots, \mathtt{open}((\theta_{N_0}^B, x_{N_0}^B), s_{N_0})\right), \end{aligned} \tag{B48}$$

leading to the state

$$\begin{aligned} \rho^{(2)} = \frac{1}{2^{N_0}} \sum_{x^B,\theta^B,r} |\beta^{x^B,\theta^B,r}|^2 |\theta^B\rangle\langle\theta^B|_{\Theta^B} |x^B\rangle\langle x^B|_{X^B} |r\rangle\langle r|_R |\phi^{x^B,\theta^B,r}\rangle\langle\phi^{x^B,\theta^B,r}|_A \\ \bigotimes_{i\in I} \left(\frac{1}{2^{n_s}} \sum_{s_i} |\mathtt{com}((\theta_i^B, x_i^B), s_i, r)\rangle\langle\mathtt{com}((\theta_i^B, x_i^B), s_i, r)|_{\mathrm{COM}_i} \otimes |\mathtt{open}((\theta_i^B, x_i^B), s_i)\rangle\langle\mathtt{open}((\theta_i^B, x_i^B), s_i)|_{\mathrm{OPEN}_i} \right). \end{aligned} \tag{B49}$$

Let $J \subseteq I$, we now want to use the hiding property of the commitment scheme to approximate the state (B49) to one where the values of com and open_J do not provide any information about $\theta_{\bar{J}}^B$. First, we proceed to rewrite the expression for the COM_i and OPEN_i subsystems by grouping the individual values of com_i

$$\begin{aligned} &\frac{1}{2^{n_s}} \sum_{s_i} |\mathtt{com}((\theta_i^B, x_i^B), s_i, r)\rangle\langle\mathtt{com}((\theta_i^B, x_i^B), s_i, r)|_{\mathrm{COM}_i} |\mathtt{open}((\theta_i^B, x_i^B), s_i)\rangle\langle\mathtt{open}((\theta_i^B, x_i^B), s_i)|_{\mathrm{OPEN}_i} \\ &\quad = \sum_{\substack{\mathrm{com}_i \\ \in\{0,1\}^{n_c}}} P_{\mathrm{com}}^{\theta_i,x_i,r}(\mathrm{com}_i) |\mathrm{com}_i\rangle\langle\mathrm{com}_i|_{\mathrm{COM}_i} \sum_{\substack{\mathrm{open}_i \\ \in \mathcal{C}_r(\mathrm{com}_i)}} P_{\mathrm{open}}^{\theta_i,x_i,r}(\mathrm{com}_i, \mathrm{open}_i) |\mathrm{open}_i\rangle\langle\mathrm{open}_i|_{\mathrm{OPEN}_i} \\ &\quad = \sigma_{\mathrm{COM}_i,\mathrm{OPEN}_i}^{\theta_i,x_i,r}, \end{aligned} \tag{B50}$$

where $P_{\mathrm{com}}^{\theta_i,x_i,r}$ is the respective distribution for com_i for uniformly sampled s_i, which depends on the commitment scheme used, and $\mathcal{C}_r(\mathrm{com}_i)$ is the set of strings open_i that satisfy $\mathtt{ver}(\mathrm{com}_i, \mathrm{open}_i, r) \neq \perp$. Substituting Eq. (B50) into Eq. (B49) and tracing out $\mathrm{OPEN}_{\bar{J}}$ and R results in

$$\rho^{(2)} = \frac{1}{2^{N_0}} \sum_{x^B,\theta^B,r} |\beta^{x^B,\theta^B,r}|^2 |\theta^B\rangle\langle\theta^B|_{\Theta^B} |x^B\rangle\langle x^B|_{X^B} |\phi^{x^B,\theta^B,r}\rangle\langle\phi^{x^B,\theta^B,r}|_A \underbrace{\bigotimes_{i\in J} \sigma_{\mathrm{COM}_i,\mathrm{OPEN}_i}^{\theta_i,x_i,r}}_{\sigma_{\mathrm{COM}_J,\mathrm{OPEN}_J}^{\theta_J,x_J,r}} \underbrace{\bigotimes_{i\in\bar{J}} \sigma_{\mathrm{COM}_i}^{\theta_i,x_i,r}}_{\sigma_{\mathrm{COM}_{\bar{J}}}^{\theta_{\bar{J}},x_{\bar{J}},r}}. \tag{B51}$$

The hiding property of the commitment scheme states that for any fixed r, the distributions $P_{\mathrm{com}}^{\theta_i,x_i,r}$ are computationally indistinguishable among themselves. Let $P_{\mathrm{com}}^r = P_{\mathrm{com}}^{0,0,r}$, then

$$\sigma_{\mathrm{COM}_i}^{\theta_i,x_i,r} \approx^{(c)} \tilde{\sigma}_{\mathrm{COM}_i}^r, \tag{B52}$$

with

$$\tilde{\sigma}_{\mathrm{COM}_i}^r = \sum_{\substack{\mathrm{com}_i \\ \in\{0,1\}^{n_c}}} P_{\mathrm{com}}^r(\mathrm{com}_i) |\mathrm{com}_i\rangle\langle\mathrm{com}_i|_{\mathrm{COM}_i}. \tag{B53}$$

Applying Eq. (B52) to the $\bar{J}$ subset in Eq. (B51), and from Lemma A2 (2) and (3) we get that

$$\rho_{\Theta^B,X^B,A,\mathrm{COM},\mathrm{OPEN}_J}^{(2)} \approx^{(c)} \tilde{\rho}_{\Theta^B,X^B,A,\mathrm{COM},\mathrm{OPEN}_J}^{(2)}, \tag{B54}$$

where

$$\tilde{\rho}^{(2)} = \frac{1}{2^{N_0}} \sum_{x^B,\theta^B,r} |\beta^{x^B,\theta^B,r}|^2 |\theta^B\rangle\langle\theta^B|_{\Theta^B} |x^B\rangle\langle x^B|_{X^B} |\phi^{x^B,\theta^B,r}\rangle\langle\phi^{x^B,\theta^B,r}|_A \sigma_{\mathrm{COM}_J,\mathrm{OPEN}_J}^{\theta_J,x_J,r} \tilde{\sigma}_{\mathrm{COM}_{\bar{J}}}^r. \tag{B55}$$

Note that, since both $\sigma^{\theta_J,x_J}_{\mathrm{COM}_J,\mathrm{OPEN}_J}$ and $\tilde{\sigma}_{\mathrm{COM}_{\bar{J}}}$ are independent of $\theta^B_{\bar{J}}, x^B_{\bar{J}}$, we can use Eq. (B47) with $I' = \bar{J}$ such that, after tracing the Θ^B_J, X^B subsystem, we obtain the state

$$\tilde{\rho}^{(2)}_{\Theta^B_{\bar{J}},A,\mathrm{COM},\mathrm{OPEN}_J} = \mathsf{U}_{\Theta^B_{\bar{J}}} \otimes \tilde{\rho}^{(2)}_{A,\mathrm{COM},\mathrm{OPEN}_J}. \tag{B56}$$

Finally, using Lemma A2 (1) and (2) we obtain the required result

$$\rho^{(2)}_{\Theta^B_{\bar{J}},A,\mathrm{COM},\mathrm{OPEN}_J} \approx^{(c)} \mathsf{U}_{\Theta^B_{\bar{J}}} \otimes \rho^{(2)}_{A,\mathrm{COM},\mathrm{OPEN}_J}. \tag{B57}$$

■

4. Proof of Lemma 5

In this section we present a proof of Lemma 5, which states that the string separation step of π_{QROT} does not leak any information about the random bit c to the receiver.

Lemma B6. Let $\mathcal{E}^{(I_t)} : \mathcal{H}_{A,\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C} \to \mathcal{H}_{A,\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C,\mathrm{SEP}}$ be the quantum operation used by Bob to compute the string separation information (J_0, J_1) during step (9) of the protocol. The resulting state after applying $\mathcal{E}^{(I_t)}$ to a product state of the form

$$\mathcal{E}^{(I_t)}(\rho_{A,\Theta^A_{\bar{I}_t}} \otimes \mathsf{U}_{\Theta^B_{\bar{I}_t}} \otimes \mathsf{U}_C) = \sigma_{A,\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C,\mathrm{SEP}} \tag{B58}$$

satisfies

$$\mathrm{Tr}_{\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t}}\left[\sigma_{A,\Theta^A_{\bar{I}_t},\Theta^B_{\bar{I}_t},C,\mathrm{SEP}}\right] = \sigma_A \otimes \sigma_{\mathrm{SEP}} \otimes \mathsf{U}_C. \tag{B59}$$

Proof. Let $\theta^{\mathrm{ch}} = \theta^A \oplus \theta^B$ and, for $b \in \{0,1\}$, define the sets $S_b = \{i \in \bar{I}_t | \theta^{\mathrm{ch}}_i = b\}$. Bob's operation consists on randomly choosing subsets I_0, I_1 of size N_{raw}, from S_0 and S_1, respectively, and then computing $J_0 = I_c, J_1 = I_{\bar{c}}$. Denote by $N_1 = (1-\alpha)N_0$ the size of the working set $\bar{I}_t$, so that $N_{\mathrm{raw}} = (\frac{1}{2} - \delta_2)N_1$. If the number of matching bases in $\bar{I}_t$, given by the Hamming weight $w_H(\theta^{\mathrm{ch}}_{\bar{I}_t})$, is either smaller than N_{raw} or greater than $N_1 - N_{\mathrm{raw}}$, Bob will not be able to construct either I_0 or I_1, in which case he sends an abort message to Alice independently of the value of c and Eq. (B59) is satisfied. On the other hand, we will show that whenever $N_{\mathrm{raw}} \le w_H(\theta^{\mathrm{ch}}_{\bar{I}_t}) \le N_1 - N_{\mathrm{raw}}$, the probability of choosing I_0, I_1 is the same for every I_0, I_1. Define

$$C(I_0, I_1) = \{\theta^{\mathrm{ch}} : \forall i \in I_0 \forall j \in I_1 (\theta^{\mathrm{ch}}_i = 0 \ \ \& \ \ \theta^{\mathrm{ch}}_j = 1)\}. \tag{B60}$$

Note that, because for any two pairs $(I^1_0, I^1_1), (I^2_0, I^2_1)$ the elements of $C(I^1_0, I^1_1)$ and $C(I^2_0, I^2_1)$ are related to each other through a permutation of indices, the size of the $C(I_0, I_1)$ is independent of I_0, I_1. The probability of Bob choosing I_0, I_1 is then given by

$$\begin{aligned}
P(I_0, I_1) &= \sum_{\theta^B_{\bar{I}_t} \in C(I_0,I_1)} P(I_0, I_1 | \theta^B_{\bar{I}_t}) P(\theta^B_{\bar{I}_t}) \\
&= \sum_{n=N_{\mathrm{raw}}}^{N_0 - N_{\mathrm{raw}}} \sum_{\substack{w_H(\theta^{\mathrm{ch}}_{\bar{I}_t})=n \\ \theta^{\mathrm{ch}} \in C(I_0,I_1)}} \binom{n}{N_{\mathrm{raw}}}^{-1} \binom{N_1 - n}{N_{\mathrm{raw}}}^{-1} P(\theta^{\mathrm{ch}}) \\
&= \sum_{n=N_{\mathrm{raw}}}^{N_0 - N_{\mathrm{raw}}} \sum_{\substack{w_H(\theta^{\mathrm{ch}}_{\bar{I}_t})=n \\ \theta^{\mathrm{ch}} \in C(I_0,I_1)}} \binom{n}{N_{\mathrm{raw}}}^{-1} \binom{N_1 - n}{N_{\mathrm{raw}}}^{-1} 2^{-N_1} \\
&= P^{\mathrm{SEP}},
\end{aligned} \tag{B61}$$

where the combinatorial factors come from the fact that, for each θ^{ch}, the I_0, I_1 are chosen uniformly among all available compatible combinations, and the 2^{-N_1} factor comes from the fact that both θ^A and θ^B are sampled independently and θ^B is sampled uniformly [as guaranteed by the product form Eq. (B58)], and the last equality comes from the fact that the number of elements in $C(I_0, I_1)$ is constant, and hence the number of terms in the summation is the same for every (I_0, I_1). Importantly, note that P^{SEP} is independent of (I_0, I_1). To obtain Eq. (B59) we start by computing

$$\begin{aligned}
\sigma_{\mathrm{SEP},\Theta^B_{\bar{I}_t},C} &= \mathcal{E}^{(I_t,\Theta^A_{\bar{I}_t})}(\mathsf{U}_{\Theta^B_{\bar{I}_t}} \otimes \mathsf{U}_C) \\
&= \frac{1}{2^{N_1}} \sum_{\theta^B_{\bar{I}_t}} |\theta^B_{\bar{I}_t}\rangle\langle\theta^B_{\bar{I}_t}|_{\Theta^B_{\bar{I}_t}} \otimes \frac{1}{2} \sum_c |c\rangle\langle c|_C \\
&\quad \otimes \sum_{I_0,I_1} P(I_0, I_1 | \theta^B_{\bar{I}_t}) |I_c, I_{\bar{c}}\rangle\langle I_c, I_{\bar{c}}|_{\mathrm{SEP}} \\
&= \frac{1}{2} \sum_c |c\rangle\langle c|_C \otimes \sum_{I_0,I_1} P(I_0, I_1) |I_c, I_{\bar{c}}\rangle\langle I_c, I_{\bar{c}}|_{\mathrm{SEP}} \\
&\quad \otimes \sum_{\theta^B_{\bar{I}_t} \in C(I_0,I_1)} P(\theta^B_{\bar{I}_t} | I_0, I_1) |\theta^B_{\bar{I}_t}\rangle\langle\theta^B_{\bar{I}_t}|_{\Theta^B_{\bar{I}_t}},
\end{aligned} \tag{B62}$$

where the sum in SEP goes over all possible I_0, I_1 given I_t. Tracing out $\Theta^B_{\bar{I}_t}$ and using Eq. (B61) we obtain

$$\begin{aligned}
&\mathrm{Tr}_{\Theta^B_{\bar{I}_t}}\left[\sigma_{\mathrm{SEP},\Theta^B_{\bar{I}_t},C}\right] \\
&= \frac{1}{2} \sum_c |c\rangle\langle c|_C \otimes \sum_{I_0,I_1} P^{\mathrm{SEP}} |I_c, I_{\bar{c}}\rangle\langle I_c, I_{\bar{c}}|_{\mathrm{SEP}} \\
&= \frac{1}{2} \sum_c |c\rangle\langle c|_C \otimes \sum_{I_0,I_1} P^{\mathrm{SEP}} |I_0, I_1\rangle\langle I_0, I_1|_{\mathrm{SEP}} \\
&= \mathsf{U}^{I_t}_{\mathrm{SEP}} \otimes \mathsf{U}_C.
\end{aligned} \tag{B63}$$

■

5. Proof of Lemma 6

In this section we present a proof of Lemma 6, introduced in Sec. IV as part of the security analysis against a dishonest receiver. Recall that the transcript of the protocol $\vec{\tau} = (x^A_{I_t}, \theta^A, r, \mathrm{com}, I_t, I_s, \mathrm{open}_{I_t}, r)$ is defined to consist of all classical information (with the exception of her measurement outcomes) that Alice has access up to step (8) of the protocol.

Lemma B7. Assuming Alice follows the protocol, let X^A, B denote the systems of Alice measurement outcomes and Bob's laboratory at the end of step (9) of the protocol, and let $\rho_{X^A,B}$ be the state of the joint system at that point. There exists a state $\tilde{\rho}_{X^A,B}$, such as

(1) The conditioned states $\tilde{\rho}_{X^A,B}(\vec{\tau}, J_0, J_1)$ satisfy

$$H_{\min}(X^A_{J_0}|X^A_{J_1}B)_{\tilde{\rho}(\vec{\tau},J_0,J_1)} + H_{\min}(X^A_{J_1}|X^A_{J_0}B)_{\tilde{\rho}(\vec{\tau},J_0,J_1)} \geq 2N_{\mathrm{raw}}\left(\frac{1}{2} - \frac{2\delta_2}{1-2\delta_2} - h\left(\frac{p_{\max}+\delta_1}{\frac{1}{2}-\delta_2}\right)\right); \tag{B64}$$

(2) $\rho_{X^A,B} \approx_\varepsilon \tilde{\rho}_{X^A,B}$, with

$$\varepsilon = \left(2\left(e^{-\frac{1}{2}\alpha(1-\alpha)^2 N_0 \delta_1^2} + e^{-\frac{1}{2}\left(\frac{1}{2}-\delta_2\right)\alpha N_0 \delta_1^2}\right)\right)^{\frac{1}{2}} + e^{-D_{\mathrm{KL}}\left(\frac{1}{2}-\delta_2|\frac{1}{2}\right)(1-\alpha)N_0} + \varepsilon_{\mathrm{bind}}(k), \tag{B65}$$

where $h(\cdot)$ and $D_{\mathrm{KL}}(\cdot\,|\,\cdot)$ denote the binary entropy and the binary relative entropy functions, respectively, and $\varepsilon_{\mathrm{bind}}(k)$ is a negligible function given by the binding property of the commitment scheme.

Proof. We proceed by tracking the properties of Alice's and Bob's shared state as the protocol develops in order to bound the conditional min-entropy of Alice's measurement outcomes given the information the Bob gains during the protocol, then we use Lemma A4 to obtain the desired result. Let $\rho^{\mathrm{rand}}_{\Theta^A,T,R}$ denote the quantum state associated to the systems holding Alice's basis choice θ^A, test subset I_t, and the value of r used in the commit and open phase, which we can treat as if they are sampled at the beginning of the protocol since their distribution is fixed, and is given by

$$\rho^{\mathrm{rand}}_{\Theta^A,T,R} = \frac{1}{2^{N_0}}\sum_{\theta^A}|\theta^A\rangle\langle\theta^A|_{\Theta^A} \otimes \frac{1}{|\mathcal{T}(\alpha N_0, I)|}\sum_{I_t}|I_t\rangle\langle I_t|_T \otimes \frac{1}{2^{n_r}}\sum_r |r\rangle\langle r|_R, \tag{B66}$$

where $\mathcal{T}(\alpha N_0, I)$ denotes the set of subsets of $I = \{1, 2, \ldots, N_0\}$ with αN_0 elements. Let $S_{\mathrm{bind}}(k)$ be the set of all $r \in \{0,1\}^{n_r(k)}$ for which there exists a tuple $(\mathrm{com}, \mathrm{open}_1, \mathrm{open}_2)$ such that

$$\perp \neq \mathtt{ver}(\mathrm{com}, \mathrm{open}_1, r) \neq \mathtt{ver}(\mathrm{com}, \mathrm{open}_2, r) \neq \perp. \tag{B67}$$

From the binding property of the commitment scheme we know that there exists a negligible function $\varepsilon_{\mathrm{bind}}(k)$ such that, for a commitment security parameter k it holds that

$$\frac{S_{\mathrm{bind}}(k)}{2^{n_r(k)}} = \varepsilon_{\mathrm{bind}}(k), \tag{B68}$$

and hence the state

$$\tilde{\rho}^{\mathrm{rand}}_{\Theta^A,T,R} = \frac{1}{2^{N_0}}\sum_{\theta^A}|\theta^A\rangle\langle\theta^A|_{\Theta^A} \otimes \frac{1}{|\mathcal{T}(\alpha N_0, I)|}\sum_{I_t}|I_t\rangle\langle I_t|_T \otimes P_R \sum_{r\in\bar{S}_{\mathrm{bind}}}|r\rangle\langle r|_R, \tag{B69}$$

where $P_R = \frac{1}{2^{n_r} - |S_{\mathrm{bind}}|}$, satisfies

$$\rho^{\mathrm{rand}}_{\Theta^A,T,R} \approx_{\varepsilon_{\mathrm{bind}}(k)} \tilde{\rho}^{\mathrm{rand}}_{\Theta^A,T,R}. \tag{B70}$$

In other words, the state of the system holding the value of the variable r is indistinguishable to one where the commitment scheme is perfectly binding (for all com strings, there is at most one open string that passes verification).

Additionally, the state of the shared resource system as after Bob receives his shares at the beginning of the protocol is given by

$$\rho^{(0)} = \tilde{\rho}^{\mathrm{rand}}_{\Theta^A,T,R} \otimes |\psi^{(0)}\rangle\langle\psi^{(0)}||\psi^{(0)}\rangle = \frac{1}{\sqrt{2^{N_0}}}\sum_x |x\rangle_{\Phi^A}|x\rangle_{\Phi^B}. \tag{B71}$$

Since the measurement on Alice subsystem is performed independently from Bob's actions, we can equivalently consider a version of the protocol in which Alice does not measure her side of the shared resource state until it is needed to perform the check at step (7) (for the indices in I_t) and the computation of the syndromes at step (10) (for the remaining indices).

We now turn our attention to step (4), when Bob computes and sends his commitment strings after receiving the value of r. Denote by B_0 the system containing all of Bob's laboratory at the beginning of the protocol, and let U_1 be the transformation that Bob performs on his system to produce the commitments, which has the general form

$$U_1|r\rangle_R|x\rangle_{\Phi^B}|0\rangle_{B_0} = \sum_{\mathrm{com}}\alpha^{r,x,\mathrm{com}}|\mathrm{com}\rangle_{\mathrm{COM}}|\phi^{r,x,\mathrm{com}}\rangle_{B_1}, \tag{B72}$$

where $\mathcal{H}_R \otimes \mathcal{H}_{\Phi^B} \otimes \mathcal{H}_{B_0} = \mathcal{H}_{\mathrm{COM}} \otimes \mathcal{H}_{B_1}$, and $\mathrm{com} = (\mathrm{com}_1, \mathrm{com}_2, \ldots, \mathrm{com}_{N_0})$ with $\mathrm{com}_i \in \{0,1\}^{n_c(k)}$. Bob then proceeds to send the COM system to Alice, who measures it in the computational basis. The joint shared state as Bob sends the commitment information is

$$\rho^{(1)} = \frac{1}{2^{N_0}|\mathcal{T}(\alpha N_0, I)|} \sum_{\theta^A} |\theta^A\rangle\langle\theta^A|_{\Theta^A} \sum_{I_t} |I_t\rangle\langle I_t|_T \sum_{\substack{r\in\bar{S}_{\mathrm{bind}} \\ \mathrm{com}}} P^r_{\mathrm{com}} |\mathrm{com}\rangle\langle\mathrm{com}|_{\mathrm{COM}} |\eta^{r,\mathrm{com}}\rangle\langle\eta^{r,\mathrm{com}}|_{\Phi^A B_1}, \tag{B73}$$

where

$$\begin{aligned} P^r_{\mathrm{com}} &= \frac{P_R}{2^{N_0}} \sum_x |\alpha^{r,x,\mathrm{com}}|^2 \\ |\eta^{r,\mathrm{com}}\rangle_{\Phi^A B_1} &= \sum_x \underbrace{\sqrt{\frac{P_R}{2^{N_0}}} (P^r_{\mathrm{com}})^{-\frac{1}{2}} \alpha^{r,x,\mathrm{com}}}_{\beta^{r,x,\mathrm{com}}} |x\rangle_{\Phi^A} |\phi^{r,x,\mathrm{com}}\rangle_{B_1}. \end{aligned} \tag{B74}$$

We intend to use Lemma B1 to bound the form of the shared state after the parameter estimation step, and then Lemma B3 to bound the amount of correlation between Alice's measurement outcomes on the system Φ^A and Bob's system. For that, we first need to associate Bob's commitments with their corresponding committed strings x^B and θ^B. For an arbitrary dishonest Bob the strings that Alice received are not guaranteed to be outcomes of the com function and may not have an associated preimage. Consider now the functions $x^B_i(r, \mathrm{com}), \theta^B_i(r, \mathrm{com}) : \{0,1\}^{n_r} \times \{0,1\}^{n_c} \to \{0,1\}$ defined as follows:

$$x^B_i(r, \mathrm{com}) = \begin{cases} x, & \text{if } \mathrm{com}_i = \mathtt{com}((\theta, x), s, r) \text{ for some } \theta \in \{0,1\}, s \in \{0,1\}^{n_s}, \\ 0, & \text{otherwise,} \end{cases} \tag{B75}$$

$$\theta^B_i(r, \mathrm{com}) = \begin{cases} \theta, & \text{if } \mathrm{com}_i = \mathtt{com}((\theta, x), s, r) \text{ for some } x \in \{0,1\}, s \in \{0,1\}^{n_s}, \\ 0, & \text{otherwise,} \end{cases} \tag{B76}$$

and denote

$$x^B(r, \mathrm{com}) = (x^B_i(r, \mathrm{com}))_i, \quad \theta^B(r, \mathrm{com}) = (\theta^B_i(r, \mathrm{com}))_i. \tag{B77}$$

We know the above functions are well defined for all $r \in \bar{S}_{\mathrm{bind}}$ because, by definition of S_{bind}, for each possible value of com_i, there is at most a single opening that passes verification. For any $\theta^B \in \{0,1\}^{N_0}$ we can write the state $|\eta^{r,\mathrm{com}}\rangle_{\Phi^A B_1}$ in the θ^B basis of Φ^A as

$$|\eta^{r,\mathrm{com}}\rangle_{\Phi^A B_1} = \sum_x \underbrace{\sum_{x'} \beta^{r,x,\mathrm{com}} \langle x, \theta^B | x'\rangle |\phi^{r,x',\mathrm{com}}\rangle_{B_1}}_{\beta^{r,x,\theta^B,\mathrm{com}} |\phi^{r,x,\theta^B,\mathrm{com}}\rangle} |x, \theta^B\rangle_{\Phi^A}. \tag{B78}$$

Recall that $I_s(\theta^A, \theta^B, I_t) = \{i \in I_t : \theta^A_i \oplus \theta^B_i = 0\}$. From Lemma B1 we know that there exists a state

$$\tilde{\rho}^{(1)} = \frac{1}{2^{N_0}|\mathcal{T}(\alpha N_0, I)|} \sum_{\substack{r\in\bar{S}_{\mathrm{bind}} \\ \mathrm{com}}} P^r_{\mathrm{com}} |\mathrm{com}\rangle\langle\mathrm{com}|_{\mathrm{COM}} \sum_{\theta^A, I_t} |\theta^A\rangle\langle\theta^A|_{\Theta^A} |I_t\rangle\langle I_t|_T |\eta^{r,\mathrm{com},I_s,\bar{I}_t}\rangle\langle\eta^{r,\mathrm{com},I_s,\bar{I}_t}|_{\Phi^A B_1}, \tag{B79}$$

where the $|\eta^{r,\mathrm{com},I_s,\bar{I}_t}\rangle$ have the form

$$|\eta^{r,\mathrm{com},I_s,\bar{I}_t}\rangle_{\Phi^A B_1} = \sum_{\substack{x\in \\ B(\theta^B,r,\mathrm{com},I_s,I_t)}} \tilde{\beta}^{x,\theta^B,r,\mathrm{com},I_s,I_t} |x, \theta^B(r,\mathrm{com})\rangle_{\Phi^A} |\phi^{x,\theta^B,r,\mathrm{com},I_s,I_t}\rangle_{B_1} \tag{B80}$$

$$B(r, \mathrm{com}, I_s, I_t) = \{x : |r_H(x_{I_s} \oplus x^B_{I_s}(r,\mathrm{com})) - r_H(x_{\bar{I}_t} \oplus x^B_{\bar{I}_t}(r,\mathrm{com}))| \le \delta_1\}, \tag{B81}$$

such that

$$D(\rho^{(1)},\tilde{\rho}^{(1)}) \leq \sqrt{2}\left(e^{-\frac{1}{2}\alpha(1-\alpha)^2 N_0\delta_1^2} + e^{-\frac{1}{2}\left(\frac{1}{2}-\delta_2\right)\alpha N_0\delta_1^2}\right)^{\frac{1}{2}}. \tag{B82}$$

We are ready now proceed to step (6) of the protocol, in which Bob sends the string $\text{open}_{I_t} = (\text{open}_i)_{i\in I_t}$, which is expected to contain the opening information for all the commitments $\text{com}_i, i \in I_t$.

$$U_{\text{open}}|\phi^{x,r,\text{com},I_s,I_t}\rangle_{B_1} = \sum_{\text{open}_{I_t}} \alpha^{x,r,\text{com},I_s,I_t,\text{open}_{I_t}}|\phi^{x,r,\text{com},I_s,I_t,\text{open}_{I_t}}\rangle_{B_2}|\text{open}_{I_t}\rangle_{\text{OPEN}_{I_t}}, \tag{B83}$$

where $\mathcal{H}_{B_1} = \mathcal{H}_{B_2}\otimes\mathcal{H}_{\text{OPEN}_{I_t}}$. Such that

$$U_{\text{open}}|\eta^{r,\text{com},I_s,\bar{I}_t}\rangle_{\Phi^A B_1} = \sum_{\text{open}_{I_t}}\sum_{x\in B}\tilde{\beta}^{x,r,\text{com},I_s,I_t}\alpha^{x,r,\text{com},I_s,I_t,\text{open}_{I_t}}|x,\theta^B(r,\text{com})\rangle_{\Phi^A}\otimes|\phi^{x,r,\text{com},I_s,I_t,\text{open}_{I_t}}\rangle_{B_2}|\text{open}_{I_t}\rangle_{\text{OPEN}_{I_t}}. \tag{B84}$$

During step (7), after receiving the opening information and measuring the OPEN system in the computational basis, she aborts the protocol unless $\texttt{ver}(\text{com}_i,\text{open}_i,r)\neq\perp$ for all $i\in I_t$. Let $H(r,I_t)$ be the set of strings com for which Alice's first check *can* be passed. From the binding property of the commitment scheme, we know that, for any $r\in\bar{S}_{\text{bind}}$, if $\text{com}\in H(r,I_t)$ there is only one $\text{open}'(r,\text{com},I_t)$ for which $\texttt{ver}(\text{com}_i,\text{open}'_i,r)\neq\perp$ for all $i\in I_t$. Because the protocol aborts if Alice's test is not passed, the state of the joint system after Alice performs this check is given by (note that from here, by removing the mixture over all opens, we are reducing the overall trace of the system). Effectively, we are keeping only the runs of the protocol that did not abort in the commitment check part of step (7). The amount for which the trace is reduced is given by the sum of the $|\alpha^{x,r,\text{com},I_s,I_t,\text{open}_{I_t}}|^2$ over the values of $\text{open}\neq\text{open}'(\text{com},I_t)$ or for which $I_s < N_{\text{check}}$):

$$\begin{aligned}\tilde{\rho}^{(2)} = \frac{1}{2^{N_0}|\mathcal{T}(\alpha N_0,I)|}\sum_{I_t,\theta^A}|I_t\rangle\langle I_t|_T \sum_{\substack{r\in\bar{S}_{\text{bind}}\\ \text{com}\in H(r,I_t)}} P^r_{\text{com}}P^r_{\text{open}'}|\text{com}\rangle\langle\text{com}|_{\text{COM}}|\text{open}'\rangle\langle\text{open}'|_{\text{OPEN}_{I_t}}\\ \otimes|\theta^A\rangle\langle\theta^A|_{\Theta^A}|\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t}\rangle\langle\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t}|_{\Phi^A B_2},\end{aligned} \tag{B85}$$

with

$$P^r_{\text{open}'} = \sum_{\substack{x\in\\ B(r,\text{com},I_s,I_t)}}|\tilde{\beta}^{x,r,\text{com},I_s,I_t}\alpha^{x,r,\text{com},I_s,I_t,\text{open}'}|^2$$

and

$$|\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t}\rangle = \sum_{\substack{x\in\\ B(r,\text{com},I_s,I_t)}}\left(\underbrace{(P^r_{\text{open}'})^{-\frac{1}{2}}\tilde{\beta}^{x,r,\text{com},I_s,I_t}\alpha^{x,r,\text{com},I_s,I_t,\text{open}'}|\phi^{x,r,\text{com},I_s,I_t,\text{open}'}\rangle_{B_2}}_{\gamma^{x,r,\text{com},I_s,I_t}|\tilde{\phi}^{x,r,\text{com},I_s,I_t}\rangle}\otimes|x,\theta^B(r,\text{com})\rangle_{\Phi^A}\right). \tag{B86}$$

Alice then proceeds to measure her part of the state. Let us divide her measurement in two parts: the measurement of the qubits in I_t, and the measurement of the reminder qubits. For the first part, the action of measuring the subsystem $\Phi^A_{I_t}$ in a state $|\tilde{\eta}^{\text{com},I_s,\bar{I}_t}\rangle$ and in the $\theta^A_{I_t}$ basis is

$$\begin{aligned}|\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t}\rangle\langle\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t}| &\to \sum_{x^A_{I_t}}|x^A_{I_t},\theta^A_{I_t}\rangle\langle x^A_{I_t},\theta^A_{I_t}|\left(|\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t}\rangle\langle\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t}|\right)|x^A_{I_t},\theta^A_{I_t}\rangle\langle x^A_{I_t},\theta^A_{I_t}|\\ &= \sum_{x^A_{I_t}}P^r_{x^A_{I_t}}|x^A_{I_t},\theta^A_{I_t}\rangle\langle x^A_{I_t},\theta^A_{I_t}|_{\Phi^A_{I_t}}|\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t,\theta^A,x^A_{I_t}}\rangle\langle\tilde{\eta}^{r,\text{com},I_s,\bar{I}_t,\theta^A,x^A_{I_t}}|_{\Phi^A_{\bar{I}_t}B_2},\end{aligned} \tag{B87}$$

where

$$P^r_{x^A_{I_t}} = \sum_{x\in B} \left|\langle x_{I_t}, \theta^B_{I_t}(r,\mathrm{com})|x^A_{I_t}, \theta^A_{I_t}\rangle \gamma^{x,r,\mathrm{com},I_s,I_t}\right|^2 \tag{B88}$$

and

$$|\tilde{\eta}^{r,\mathrm{com},I_s,\bar{I}_t,x^A_{I_t},\theta^A}\rangle = \sum_{x\in B}(P^r_{x^A_{I_t}})^{-\frac{1}{2}} \underbrace{\langle x_{I_s}, \theta^B_{I_s}|x^A_{I_s}, \theta^B_{I_s}\rangle}_{\delta(x_{I_s},x^A_{I_s})} \langle x_{I_t\setminus I_s}, \theta^B_{I_t\setminus I_s}|x^A_{I_t\setminus I_s}, \bar{\theta}^B_{I_t\setminus I_s}\rangle \gamma^{x,\theta^B,r,\mathrm{com},I_s,I_t} |\tilde{\phi}^{x,\theta^A,r,\mathrm{com},I_t}\rangle_{B_2} |x_{\bar{I}_t}, \theta^B_{\bar{I}_t}\rangle_{\Phi^A_{\bar{I}_t}}, \tag{B89}$$

where in the last expression, and going forward, we omit the explicit dependence of both x^B and θ^B on r, com. By defining

$$G(x^A_{I_s}, r, \mathrm{com}) = \{x_{\bar{I}_t} : |r_H(x^A_{I_s} \oplus x^B_{I_s}) - r_H(x_{\bar{I}_t} \oplus x^B_{\bar{I}_t})| \le \delta_1\}, \tag{B90}$$

we can rewrite

$$|\tilde{\eta}^{\mathrm{com},I_s,\bar{I}_t,x^A_{I_t},\theta^A}\rangle = \sum_{x_{\bar{I}_t}\in G} \underbrace{\sum_{x_{I_t\setminus I_s}} (P^r_{x^A_{I_t}})^{-\frac{1}{2}} \langle x_{I_t\setminus I_s}, \theta^B_{I_t\setminus I_s}|x^A_{I_s}, \theta^B_{I_s}\rangle \gamma^{x_{\bar{I}_t},x^A_{I_t},r,\mathrm{com},I_s,I_t} |\tilde{\phi}^{x_{\bar{I}_t},x^A_{I_t},r,\mathrm{com},I_s,I_t}\rangle_{B_2}}_{\tilde{\gamma}^{x_{\bar{I}_t},r,\mathrm{com},I_s,I_t}|\tilde{\phi}^{x_{\bar{I}_t},r,\mathrm{com},I_s,I_t,\theta^A}\rangle} \otimes |x_{\bar{I}_t}, \theta^B_{\bar{I}_t}\rangle_{\Phi^A_{\bar{I}_t}}. \tag{B91}$$

After performing the measurement, Alice aborts the protocol whenever $r_H(x^A_{I_s} \oplus x^B_{I_s}(r,\mathrm{com})) > p_{\max}$. The state of the shared system after this check is (tracing out the T, COM, OPEN subsystems)

$$\begin{aligned}\tilde{\rho}^{(3)} = \frac{1}{2^{N_0}|\mathcal{T}(\alpha N_0, I)|} \sum_{I_t} \sum_{\substack{r\in\bar{S}_{\mathrm{bind}} \\ \mathrm{com}\in H(r,I_t)}} P^r_{\mathrm{com}} P^r_{\mathrm{open}'} \sum_{\theta^A} |\theta^A\rangle\langle\theta^A|_{\Theta^A} \\ \otimes \sum_{\substack{x^A_{I_t}\in \\ J_{p_{\max}}}} P^r_{x^A_{I_t}} |x^A_{I_t}, \theta^A_{I_t}\rangle\langle x^A_{I_t}, \theta^A_{I_t}|_{\Phi^A_{I_t}} |\tilde{\eta}^{r,\mathrm{com},I_s,\bar{I}_t,x^A_{I_t},\theta^A}\rangle\langle\tilde{\eta}^{r,\mathrm{com},I_s,\bar{I}_t,x^A_{I_t},\theta^A}|_{\Phi^A_{\bar{I}_t}B_2},\end{aligned} \tag{B92}$$

where

$$J_{p_{\max}} = \{x^A_{I_t} : r_H(x^A_{I_s} \oplus x^B_{I_s}) \le p_{\max}\}. \tag{B93}$$

Before proceeding, it will be useful to approximate the above state to a state where the number of mismatching bases in $\bar{I}_t$ is "high enough." More precisely, this means approximating $\tilde{\rho}^{(3)}$ to a state for which the sum over θ^A runs explicitly over strings $\theta^A \in K_{\mathrm{com}} = \{\theta^A : w_H(\theta^{\mathrm{ch}}_{\bar{I}_t}) \ge N_{\mathrm{raw}}\}$. Let

$$\begin{aligned}\tilde{\tilde{\rho}}^{(3)} = \frac{1}{2^{N_0}|\mathcal{T}(\alpha N_0, I)|} \sum_{I_t} \sum_{\substack{r\in\bar{S}_{\mathrm{bind}} \\ \mathrm{com}\in H(r,I_t)}} P^r_{\mathrm{com}} P^r_{\mathrm{open}'} a(\mathrm{com}) \sum_{\theta^A\in K_{\mathrm{com}}} |\theta^A\rangle\langle\theta^A|_{\Theta^A} \\ \otimes \sum_{\substack{x^A_{I_t}\in \\ J_{p_{\max}}}} P^r_{x^A_{I_t}} |x^A_{I_t}, \theta^A_{I_t}\rangle\langle x^A_{I_t}, \theta^A_{I_t}|_{\Phi^A_{I_t}} |\tilde{\eta}^{r,\mathrm{com},I_s,\bar{I}_t,x^A_{I_t}}\rangle\langle\tilde{\eta}^{r,\mathrm{com},I_s,\bar{I}_t,x^A_{I_t}}|_{\Phi^A_{\bar{I}_t}B_2},\end{aligned} \tag{B94}$$

with $a(\mathrm{com}) = 2^{N_0}/|K_{\mathrm{com}}|$. The distance between $\tilde{\rho}^{(3)}$ and $\tilde{\tilde{\rho}}^{(3)}$ is bounded by the probability of a uniformly chosen θ^A not being in K_{com}. Using the Chernoff-Hoeffding bound we get

$$D(\tilde{\rho}^{(3)}, \tilde{\tilde{\rho}}^{(3)}) \le e^{-D_{\mathrm{KL}}(\frac{1}{2}-\delta_2|\frac{1}{2})(1-\alpha)N_0}, \tag{B95}$$

where $D_{\mathrm{KL}}(\frac{1}{2}-\delta_2 \mid \frac{1}{2})$ represents the relative entropy between the binary distributions defined by the respective probabilities $p_1 = \frac{1}{2}-\delta_2$ and $p_2 = \frac{1}{2}$.

During step (8), Alice sends the Θ^A system to Bob, who then computes J_0, J_1 (in the actual protocol, Alice sends only $\Theta^A_{\bar{I}_t}$, but to simplify the expressions we can assume, without loss of generality, that she sends the whole register Θ^A). To simplify the list of dependencies, denote the transcript of the protocol up until step (8) as $\vec{\tau} = (x^A_{I_t}, \theta^A, r, \mathrm{com}, I_t, I_s, \mathrm{open}_{I_t})$. Keep in mind that, although $\vec{\tau}$ consists of seven quantities, I_s and open_{I_t} are completely defined by the other five. In the remaining of the proof, unless noted otherwise, the sums over $\vec{\tau}$ run over the values of its variables as shown in Eq. (B94). By defining

$$P_{\vec{\tau}} = \frac{P^r_{\mathrm{com}} P^r_{\mathrm{open}'} P^r_{x^A_{I_t}} a(\mathrm{com})}{2^{N_0} |\mathcal{T}(\alpha N_0, I)|}, \tag{B96}$$

we can write

$$\tilde{\tilde{\rho}}^{(3)} = \sum_{\vec{\tau}} P_{\vec{\tau}} |\theta^A\rangle\langle\theta^A|_{\Theta^A} |x^A_{I_t}, \theta^A_{I_t}\rangle\langle x^A_{I_t}, \theta^A_{I_t}|_{\Phi^A_{\bar{I}_t}} |\tilde{\eta}^{r,\mathrm{com},I_s,\bar{I}_t,x^A_{\bar{I}_t}}\rangle\langle\tilde{\eta}^{r,\mathrm{com},I_s,\bar{I}_t,x^A_{\bar{I}_t}}|_{\Phi^A_{\bar{I}_t} B_2}. \tag{B97}$$

During step (9), after receiving θ^A, Bob sends the SEP system, containing the (classical) string separation information J_0, J_1 to Alice. By following the same treatment as in steps (4) and (6), let U_{sep} be the operation that Bob performs on the B_2 system to compute the information to be sent to Alice in the SEP system:

$$U_{\mathrm{sep}} |\tilde{\phi}^{x_{\bar{I}_t}, r, \mathrm{com}, I_s, I_t, \theta^A}\rangle_{B_2} |\theta^A\rangle_{\Theta^A} = \sum_{J_0, J_1} \alpha^{x_{\bar{I}_t}, \vec{\tau}, J_0, J_1} |\phi^{x_{\bar{I}_t}, \vec{\tau}, J_0, J_1}\rangle_{B_3} |J_0, J_1\rangle_{\mathrm{SEP}}, \tag{B98}$$

where $\mathcal{H}_{B_2} \otimes \mathcal{H}_{\Theta^A} = \mathcal{H}_{B_3} \otimes \mathcal{H}_{\mathrm{SEP}}$ and the summation over J_0, J_1 goes over all possible values compatible with I_t. The state after step (9) after Alice receives the SEP system and measures in the computational basis is then given by (tracing out SEP)

$$\tilde{\rho}^{(4)} = \sum_{\substack{\vec{\tau} \\ J_0, J_1}} P_{\vec{\tau}, J_0, J_1} |x^A_{I_t}, \theta^A_{I_t}\rangle\langle x^A_{I_t}, \theta^A_{I_t}|_{\Phi^A_{I_t}} |\nu^{\vec{\tau}, J_0, J_1}\rangle\langle\nu^{\vec{\tau}, J_0, J_1}|_{\Phi^A_{\bar{I}_t} B_3}, \tag{B99}$$

where

$$\begin{aligned} P_{\vec{\tau}, J_0, J_1} &= \sum_{x_{\bar{I}_t} \in G} |\tilde{\gamma}^{x_{\bar{I}_t}, r, \mathrm{com}, I_s, I_t} \alpha^{x_{\bar{I}_t}, \vec{\tau}, J_0, J_1}|^2 \\ |\nu^{\vec{\tau}, J_0, J_1}\rangle_{\Phi^A_{\bar{I}_t} B_3} &= \sum_{x_{\bar{I}_t} \in G} \underbrace{(P_{\vec{\tau}, J_0, J_1})^{-\frac{1}{2}} \tilde{\gamma}^{x_{\bar{I}_t}, r, \mathrm{com}, I_s, I_t} \alpha^{x_{\bar{I}_t}, \vec{\tau}, J_0, J_1}}_{\beta^{x_{\bar{I}_t}, \vec{\tau}, J_0, J_1}} |x_{\bar{I}_t}, \theta^B_{\bar{I}_t}\rangle_{\Phi^A_{\bar{I}_t}} |\phi^{x_{\bar{I}_t}, \vec{\tau}, J_0, J_1}\rangle_{B_3}. \end{aligned} \tag{B100}$$

We can now consider Alice's measurement on the $\Phi^A_{\bar{I}_t}$ system. So far we have tracked the evolution of the joint state in order to describe the relationship between both parties' information. To finalize the proof we only need to keep track of the conditional min-entropy of Alice's outcomes given Bob's part of the joint system. Let $\rho_{X^A_{\bar{I}_t} B_3}(\vec{\tau}, J_0, J_1)$ be the resulting (conditioned) state after measuring the system $\Phi^A_{\bar{I}_t}$ in the $\theta^A_{\bar{I}_t}$ basis, recording the respective outcomes in the $X^A_{\bar{I}_t}$ system, and tracing out the $\Phi^A_{\bar{I}_t}$ subsystem. We can write the state of the joint system after the measurement as

$$\tilde{\rho}^{(5)} = \sum_{\substack{\vec{\tau} \\ J_0, J_1}} P_{\vec{\tau}, J_0, J_1} \tilde{\rho}_{X^A_{\bar{I}_t} B_3}(\vec{\tau}, J_0, J_1). \tag{B101}$$

Additionally, for any given J_0, J_1, denote by J_d the complement of $J_0 \cup J_1$ in $\bar{I}_t$. Following Lemma A3 (3) and (5) we know that for any $b \in \{0, 1\}$

$$\begin{aligned} H^{\varepsilon}_{\min}(X^A_{J_b} \mid X^A_{J_{\bar{b}}} B_3)_{\tilde{\rho}^{(5)}} &\geq H^{\varepsilon}_{\min}(X^A_{J_b} \mid X^A_{J_{\bar{b}}} X^A_{J_d} B_3)_{\tilde{\rho}^{(5)}} \geq \inf_{\vec{\tau}, J_0, J_1} \left\{ H^{\varepsilon}_{\min}(X^A_{J_b} \mid X^A_{J_{\bar{b}}} X^A_{J_d} B_3)_{\tilde{\rho}(\vec{\tau}, J_0, J_1)} \right\} \\ &\geq \inf_{\vec{\tau}, J_0, J_1} \left\{ \inf_{x_{J_{\bar{b},d}}} \left\{ H^{\varepsilon}_{\min}(X^A_{J_b} | B_3)_{\tilde{\rho}(x_{J_{\bar{b},d}}, \vec{\tau}, J_0, J_1)} \right\} \right\}. \end{aligned} \tag{B102}$$

We can invoke Lemma B3 to obtain an expression for the above quantity explicitly in terms of the protocol parameters N_0, α, δ_1, and δ_2. For that, we must take a small detour to define the associated mixed states $\rho^{\text{mix}}_{\Phi^A_{J_b}B_3}$ and compute their respective post-measurement entropy. First, for $b \in \{0,1\}$, we compute the reduced states

$$\rho_{\Phi^A_{J_b}B_3}(\vec{\tau}, J_0, J_1) = \text{Tr}_{\Phi^A_{J_{\bar{b},d}}}\left[|\nu^{\vec{\tau},J_0,J_1}\rangle\langle\nu^{\vec{\tau},J_0,J_1}|_{\Phi^A_{\bar{I}_t}B_3}\right] = \sum_{x_{J_{\bar{b},d}}} P^{\vec{\tau},J_0,J_1}_{x_{J_{\bar{b},d}}} |\nu^{x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1}\rangle\langle\nu^{x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1}|_{\Phi^A_{J_b}B_3}, \tag{B103}$$

with

$$\begin{aligned} P^{\vec{\tau},J_0,J_1}_{x_{J_{\bar{b},d}}} &= \sum_{\substack{x_{J_b}\in \\ B_b(x_{J_{\bar{b},d}})}} \left|\beta^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}\right|^2 \\ |\nu^{x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1}\rangle_{\Phi^A_{J_b}B_3} &= \sum_{\substack{x_{J_b}\in \\ B_b(x_{J_{\bar{b},d}})}} \underbrace{(P^{\vec{\tau},J_0,J_1}_{x_{J_{\bar{b},d}}})^{-\frac{1}{2}}\beta^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}}_{\lambda^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}} |x_{J_b},\theta^B_{J_b}\rangle_{\Phi^A_{J_b}} |\phi^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}\rangle_{B_3}, \end{aligned} \tag{B104}$$

and

$$\begin{aligned} B_b(x_{J_{\bar{b},d}}) &= \left\{x_{J_b} : x_{J_{b,\bar{b},d}} \in G(x^A_{I_s}, r, \text{com})\right\} \\ &= \left\{x_{J_b} : \left|\left(\frac{1}{2}-\delta_2\right) r_H\left(x_{J_b}\oplus x^B_{J_b}\right) + \left(\frac{1}{2}+\delta_2\right) r_H\left(x_{J_{\bar{b},d}}\oplus x^B_{J_{\bar{b},d}}\right) - r_H\left(x^A_{I_s}\oplus x^B_{I_s}\right)\right| \le \delta_1\right\}, \end{aligned} \tag{B105}$$

where the explicit dependence of B_b on $x^A_{I_s}, r, \text{com}$ has been omitted for compactness. Note that since $r_H(x^A_{I_s}\oplus x^B_{I_s}) \le p_{\max}$ the size of $B_b(x_{J_{\bar{b},d}})$ is upper bounded by

$$|B_b(x_{J_{\bar{b},d}})| \le 2^{h\left(\frac{p_{\max}+\delta_1}{\frac{1}{2}-\delta_2}\right)N_{\text{raw}}}, \tag{B106}$$

where the h stands for the binary entropy function. We can now define

$$\rho^{\text{mix}}_{\Phi^A_{J_b}B_3}(x_{J_{\bar{b},d}}, \vec{\tau}, J_0, J_1) = \sum_{\substack{x_{J_b}\in \\ B_b(x_{J_{\bar{b},d}})}} \left|\lambda^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}\right|^2 |x_{J_b},\theta^B_{J_b}\rangle\langle x_{J_b},\theta^B_{J_b}|_{\Phi^A_{J_b}} \otimes |\phi^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}\rangle\langle\phi^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}|_{B_3}. \tag{B107}$$

Measuring the above state in the $\theta^A_{J_b}$ basis, recording the results in X_{J_b} and tracing out $\Phi^A_{J_b}$ leads to

$$\rho^{\text{mix}}_{X^A_{J_b}B_3}(x_{J_{\bar{b},d}}, \vec{\tau}, J_0, J_1) = \sum_{x^A_{J_b}} \sum_{\substack{x_{J_b}\in \\ B_c(x_{J_{\bar{b},d}})}} \left|\lambda^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}\right|^2 |\langle x^A_{J_b},\theta^A_{J_b}|x_{J_b},\theta^B_{J_b}\rangle|^2 |x^A_{J_b}\rangle\langle x^A_{J_b}|_{X^A_{J_b}} \otimes |\phi^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}\rangle\langle\phi^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}|_{B_3}, \tag{B108}$$

defining $J^{0/1}_b = \{i \in J_b : \theta^{\text{ch}}_i = 0/1\}$ we can write the factors

$$\left|\langle x^A_{J_b},\theta^A_{J_b}|x_{J_b},\theta^B_{J_b}\rangle\right|^2 = \prod_{i\in J^0_b} \underbrace{\left|\langle x^A_i,\theta^A_i|x_i,\theta^B_i\rangle\right|^2}_{\delta(x^A_i,x_i)} \prod_{i\in J^1_b} \underbrace{\left|\langle x^A_i,\theta^A_i|x_i,\theta^B_i\rangle\right|^2}_{\left|\frac{1}{\sqrt{2}}\right|^2}, \tag{B109}$$

substituting in Eq. (B108) we get

$$\rho^{\mathrm{mix}}_{X^A_{J_b}B_3}(x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1)=\left(\frac{1}{2}\right)^{w_H(\theta^{\mathrm{ch}}_{J_b})}\sum_{x^A_{J_b^1}}|x^A_{J_b^1}\rangle\langle x^A_{J_b^1}|_{X^A_{J_b^1}}\otimes\sum_{\substack{x_{J_b}\in\\ B_b(x_{J_{\bar{b},d}})}}\left|\lambda^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}\right|^2|x_{I_b^0}\rangle\langle x_{I_b^0}|_{X^A_{J_b^0}}|\phi^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}\rangle\langle\phi^{x_{\bar{I}_t},\vec{\tau},J_0,J_1}|_{B_3}, \tag{B110}$$

which is a product state between the systems $X^A_{J_b^1}$ and $X^A_{J_b^0}B_3$. From Lemma A3 (1) and (2) we know that

$$\begin{aligned}H^{\varepsilon}_{\min}(X^A_{J_b}|B_3)_{\rho^{\mathrm{mix}}(x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1)}&\geq H^0_{\min}(X^A_{J_b}|B_3)_{\rho^{\mathrm{mix}}(x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1)}\\&\geq H^0_{\min}(X^A_{J_b^0}|B_3)_{\rho^{\mathrm{mix}}(x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1)}+H^0_{\min}(X^A_{J_b^1})_{\rho^{\mathrm{mix}}(x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1)}\\&\geq H^0_{\min}(X^A_{J_b^1})_{\rho^{\mathrm{mix}}(x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1)}\\&=-\log\left(\left(\frac{1}{2}\right)^{w_H(\theta^{\mathrm{ch}}_{J_b})}\right)=w_H(\theta^{\mathrm{ch}}_{J_b}).\end{aligned} \tag{B111}$$

Application of Lemma B3 together with Eqs. (B111) and (B106) leads to

$$\begin{aligned}H^{\varepsilon}_{\min}(X^A_{J_b}\mid B_3)_{\tilde{\rho}(x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1)}&\geq H^{\varepsilon}_{\min}(X^A_{J_b}\mid B_3)_{\rho^{\mathrm{mix}}(x_{J_{\bar{b},d}},\vec{\tau},J_0,J_1)}-\log\left(|B_b(x_{J_{\bar{b},d}})|\right)\\&\geq w_H(\theta^{\mathrm{ch}}_{J_b})-h\left(\frac{p_{\max}+\delta_1}{\frac{1}{2}-\delta_2}\right)N_{\mathrm{raw}}.\end{aligned} \tag{B112}$$

Note that the above expression depends only on the number of nonmatching bases θ^{ch} associated to the indices in J_b and the parameters of the protocol, which in turn makes the infimum in Eq. (B102) straightforward to compute. We can now add the respective conditional min-entropies for $X^A_{J_0}$ and $X^A_{J_1}$, which results in

$$\begin{aligned}H^{\varepsilon}_{\min}(X^A_{J_0}\mid X^A_{J_1}B_3)_{\tilde{\rho}(\vec{\tau},J_0,J_1)}+H^{\varepsilon}_{\min}(X^A_{J_1}|X^A_{J_0}B_3)_{\tilde{\rho}(\vec{\tau},J_0,J_1)}&\geq w_H(\theta^{\mathrm{ch}}_{J_0})+w_H(\theta^{\mathrm{ch}}_{J_1})-2h\left(\frac{p_{\max}+\delta_1}{\frac{1}{2}-\delta_2}\right)N_{\mathrm{raw}}\\&\geq N_{\mathrm{raw}}-2\delta_2(1-\alpha)N_0-2h\left(\frac{p_{\max}+\delta_1}{\frac{1}{2}-\delta_2}\right)N_{\mathrm{raw}}\\&\geq 2N_{\mathrm{raw}}\left(\frac{1}{2}-\frac{2\delta_2}{1-2\delta_2}-h\left(\frac{p_{\max}+\delta_1}{\frac{1}{2}-\delta_2}\right)\right).\end{aligned} \tag{B113}$$

The result follows by recalling, from Eqs. (B82) and (B95), that the real state at this point in the protocol has distance from $\tilde{\rho}^{(5)}$ bounded by

$$\varepsilon=\sqrt{2}\left(e^{-\frac{1}{2}\alpha(1-\alpha)^2N_0\delta_1^2}+e^{-\frac{1}{2}\left(\frac{1}{2}-\delta_2\right)\alpha N_0\delta_1^2}\right)^{\frac{1}{2}}+e^{-D_{\mathrm{KL}}(\frac{1}{2}-\delta_2|\frac{1}{2})(1-\alpha)N_0}+\varepsilon_{\mathrm{bind}}(k). \tag{B114}$$

■

APPENDIX C: UC SECURITY IN THE RANDOM ORACLE MODEL

Following the discussion made in Sec. IV D, we prove the composability of a specific family of weakly interactive commitment schemes in the *classical access random oracle model*, which we will refer to as ROM from here onwards. These commitments, originally proposed by Lorünser *et al.* [67], build upon the original Naor bit commitment [35] and efficiently generalize it for arbitrary k-bit string commitments without the need of error-correcting codes. A description of the commitment protocol is shown in Fig. 7, whose correctness, binding, and hiding properties, have been proven in Ref. [67]. Instead, we will thus limit ourselves to prove that the LRV commitment protocol UC-emulates the commitment functionality $\mathcal{F}_{\text{COM}}$ (defined in Fig. 8) when the hash function is modeled as an oracle $\mathcal{F}_{\text{RO}}$ which computes a random function.

Let Π_A and Π_B represent the programs for the verifier and prover, respectively, as shown in Fig. 7. Note that, for simplicity, the external inputs that trigger the start and end of the commit and reveal phases have been omitted from Fig. 7; without loss of generality, we can consider them to take the form of the respective inputs and outputs as shown in the $\mathcal{F}_{\text{COM}}$ functionality. More specifically, the commit phase starts when Π_B receives the input $(\texttt{commit}, sid, \mathbf{b})$ and ends when Π_A outputs $(\texttt{committed}, sid)$, etc. We proceed now to separate the security in two cases, in which the adversary controls Alice or Bob, respectively, as shown in Fig. 9. In order to prove security we must show that for any efficient (i.e., polynomial-time) adversary Adv with classical access to the oracle there exists a respective simulator $\mathcal{S}$ such that for any environment, which is able to send and receive inputs and outputs through the loose wires in the right and left of the diagrams, the *real-world* and *ideal-world* scenarios are indistinguishable. Denote by H the function that the random oracle computes.

a. Dishonest Bob

We construct the simulator in terms of the following subprograms:

(1) $\mathcal{F}^*_{\text{RO}}$: The same as $\mathcal{F}_{\text{RO}}$, except that it saves a list L of all the queries that have been made to the internal memory of $\mathcal{S}$.
(2) Π^*_A: The same as Π_A, except that after receiving c from Adv it runs through the current list L of queries. When it finds an $\mathbf{x}' \in L$ and $\mathbf{b} \in \{0,1\}^n$ such that

$$\mathbf{c} = G(\mathbf{x}') \oplus \sum_{i=1}^{n} b_i \cdot \mathbf{r}_i, \tag{C1}$$

it sends $(\texttt{commit}, sid, \mathbf{b})$ to $\mathcal{F}_{\text{COM}}$. If no pair $(\mathbf{b}, \mathbf{x}')$ is found, it samples uniformly a value $\mathbf{b}$ and sends $(\texttt{commit}, sid, \mathbf{b})$ to $\mathcal{F}_{\text{COM}}$. In the reveal phase, if the check is passed, it sends $(\texttt{open}, sid)$ to $\mathcal{F}_{\text{COM}}$.

Because of the binding property of the commitment protocol, the simulator may find at most one pair $(\mathbf{b}, \mathbf{x}')$ satisfying Eq. (C1) when looking through the list, except

LRV25 String commitment protocol

Parameters:

- Parties Alice (Verifier) and Bob (Prover)
- Security parameter k and message length n
- Collision-resistant hash function $H : \{0,1\}^k \to \{0,1\}^{3k+n}$
- A subroutine O, which on input a vector $\mathbf{r}_1 \in \{0,1\}^{3k+n}$, outputs a tuple of n linearly independent vectors $(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ in $\{0,1\}^{3k+n}$

Inputs:

- Bob receives the n-bit string $\mathbf{b} = (b_1, \ldots, b_n)$

(Commit phase)

1. Alice uniformly samples a $(3k+n)$-bit string $\mathbf{r}_1$ and sends it to Bob
2. Bob uniformly samples an n-bit string $\mathbf{x}$ and computes $(\mathbf{r}_1, \ldots, \mathbf{r}_n) = O(\mathbf{r}_1)$. Then, he computes $\mathbf{c} = H(\mathbf{x}) \oplus \sum_{i=1}^{n} b_i \cdot \mathbf{r}_i$ and sends $\mathbf{c}$ to Alice

(Open phase)

1. Bob sends $(\mathbf{b}, \mathbf{x})$ to Alice
2. Alice computes $(\mathbf{r}_1, \ldots, \mathbf{r}_n) = O(\mathbf{r}_1)$, then, if $\mathbf{c} = H(\mathbf{x}) \oplus \sum_{i=1}^{n} b_i \cdot \mathbf{r}_i$ outputs $\mathbf{b}$.

FIG. 7. Weakly interactive string commitment scheme based on hash functions.

Functionality $\mathcal{F}_{\text{COM}}$

Parameters:

- Parties Alice (Verifier) and Bob (Prover)
- Message length n

1. Upon receiving an input $(\texttt{commit}, sid, \mathbf{b})$ from Bob, if no value has previously been committed, output the message $(\texttt{committed}, sid)$ to Alice
2. Upon receiving the input $(\texttt{open}, sid)$ from Bob, if a value $\mathbf{b}$ has previously been committed, output the message $(\texttt{open}, sid, \mathbf{b})$ to Alice

FIG. 8. Commitment ideal functionality.

with negligible probability (this is because the probability of there existing more than one valid openings for a given value of $\mathbf{c}$ is negligible). This allows $\mathcal{S}$ to correctly extract the committed value from $\mathbf{c}$ and commit it to $\mathcal{F}_{\text{COM}}$. Note that in the case no valid opening is found from L, the simulator commits a random value to $\mathcal{F}_{\text{COM}}$. If the adversary is able to provide a valid opening pair $(\mathbf{b}, \mathbf{x})$ in the Reveal phase the two scenarios could be distinguished. However, from the preimage resistance of random oracles, an efficient adversary cannot find a valid opening from a value of $\mathbf{c}$ without having obtained it by querying the oracle, meaning that regardless of $\mathcal{S}$ committing a random value to $\mathcal{F}_{\text{COM}}$, the probability of it being opened is negligible.

b. Dishonest Alice

Similarly, we construct the simulator in terms of the following subprograms:

(1) $\mathcal{F}^*_{\text{RO}}$: The same as for the dishonest Bob case, except it may be reprogrammed on individual query-output pairs.
(2) Π^*_B: The same as Π_B, except upon receiving an input of the form $(\texttt{committed}, sid)$ from $\mathcal{F}_{\text{COM}}$, it samples uniformly the value $\mathbf{c}'$ and sends it to Adv. In the Reveal phase, upon receiving $(\texttt{open}, sid, \mathbf{b})$ from $\mathcal{F}_{\text{COM}}$, samples a random $\mathbf{x}'$ not in L, sets $\mathcal{F}^*_{\text{RO}}$ so that

$$G(\mathbf{x}') = \mathbf{c}' \oplus \sum_{i=1}^{n} b_i \cdot \mathbf{r}_i, \tag{C2}$$

and sends $(\mathbf{b}, \mathbf{x}')$ to Adv.

From the hiding property of the commitment protocol, the value $\mathbf{c}$ received by Alice during the commit phase does not give a significant advantage to an efficient adversary in finding the committed value $\mathbf{b}$ as compared to a random string. Because of this, an efficient adversary cannot distinguish if the randomly sampled $\mathbf{c}'$ corresponds to any possible committed value, except with negligible probability. During the reveal phase, the reprogramming of the oracle according to Eq. (C2) guarantees that have $\mathbf{c}'$ will be consistent with the committed values from $\mathcal{F}_{\text{COM}}$. The only difference between the real and ideal scenarios is the change in the behavior of the oracle. Because the value $\mathbf{c}'$ was sampled uniformly, the associated outcome $G(\mathbf{x}')$ as defined by Eq. (C2) is also uniformly distributed and independent on the rest of the values $G(\mathbf{x} \neq \mathbf{x}')$, resulting in both scenarios being consistent with the oracle computing a random function, and therefore indistinguishable from each other.

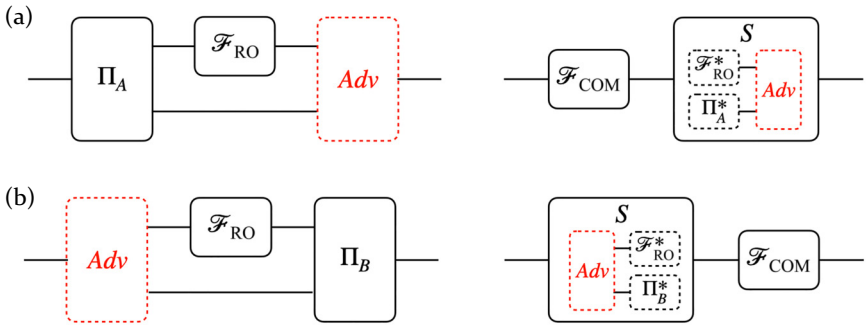


FIG. 9. Box diagrams for the execution of the protocol for (a) dishonest Bob and (b) dishonest Alice. The left sides represent the *real world* protocol interacting with an adversary Adv while the right sides represent the *ideal world* functionality interacting with the respective simulator $\mathcal{S}$.

[1] C. E. Shannon, Communication theory of secrecy systems, Bell Syst. Tech. J. **28**, 656 (1949).

[2] P. W. Shor, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (IEEE, Washington, DC, USA, 1994), pp. 124–134.

[3] V. Lyubashevsky, C. Peikert, and O. Regev, On ideal lattices and learning with errors over rings, J. ACM **60**, 1 (2013).

[4] O. Regev, On lattices, learning with errors, random linear codes, and cryptography, J. ACM **56**, 1 (2009).

[5] D. J. Bernstein and T. Lange, Post-quantum cryptography, Nature **549**, 188 (2017).

[6] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, *et al.*, Advances in quantum cryptography, Adv. Opt. Photonics **12**, 1012 (2020).

[7] S. Wiesner, Conjugate coding, ACM Sigact News **15**, 78 (1983).

[8] A. Broadbent and C. Schaffner, Quantum cryptography beyond quantum key distribution, Des. Codes Cryptogr. **78**, 351 (2016).

[9] Y. Lindell, in *Encyclopedia of Data Warehousing and Mining* (IGI Global, Hershey, PA, USA, 2005) pp. 1005–1009.

[10] Y. Lindell, Secure multiparty computation (MPC), Cryptology ePrint Archive (2020).

[11] C. Zhao, S. Zhao, M. Zhao, Z. Chen, C.-Z. Gao, H. Li, and Y.-A. Tan, Secure multi-party computation: Theory, practice and applications, Inf. Sci. **476**, 357 (2019).

[12] A. C. Yao, in *23rd Annual Symposium on Foundations of Computer Science (SFCS 1982)* (IEEE, Chicago, IL, USA, 1982), pp. 160–164.

[13] O. Goldreich, S. Micali, and A. Wigderson, in *Providing Sound Foundations for Cryptography: On the Work of Shafi Goldwasser and Silvio Micali* (Association for Computing Machinery, New York, NY, USA, 2019), pp. 307–328.

[14] I. Damgård, V. Pastro, N. Smart, and S. Zakarias, in *Advances in Cryptology–CRYPTO 2012: 32nd Annual Cryptology Conference, Santa Barbara, CA, USA, August 19–23, 2012. Proceedings* (Springer, Santa Barbara, CA, USA, 2012), pp. 643–662.

[15] S. Even, O. Goldreich, and A. Lempel, A randomized protocol for signing contracts, Commun. ACM **28**, 637 (1985).

[16] M. O. Rabin, How to exchange secrets with oblivious transfer, Technical Report, TR-81, Aiken Computation Lab, Harvard University, 1981.

[17] C. Crépeau, in *Advances in Cryptology-CRYPTO'87: Proceedings 7* (Springer, Santa Barbara, CA, USA, 1988), pp. 350–354.

[18] A. C.-C. Yao, in *27th Annual Symposium on Foundations of Computer Science (SFCS 1986)* (IEEE, Washington, DC, USA, 1986), pp. 162–167.

[19] J. Kilian, in *Proceedings of the Twentieth Annual ACM symposium on Theory of Computing* (Association for Computing Machinery (ACM), Chicago, IL, USA, 1988), pp. 20–31.

[20] M. Lemus, M. F. Ramos, P. Yadav, N. A. Silva, N. J. Muga, A. Souto, N. Paunković, P. Mateus, and A. N. Pinto, Generation and distribution of quantum oblivious keys for secure multiparty computation, Appl. Sci. **10**, 4080 (2020).

[21] C. H. Bennett, G. Brassard, C. Crépeau, and M.-H. Skubiszewska, in *Annual International Cryptology Conference* (Springer, Santa Barbara, CA, USA, 1991), pp. 351–366.

[22] R. Canetti and M. Fischlin, in *Advances in Cryptology–CRYPTO 2001: 21st Annual International Cryptology Conference, Santa Barbara, California, USA, August 19–23, 2001 Proceedings 21* (Springer, Santa Barbara, CA, USA, 2001), pp. 19–40.

[23] H.-K. Lo and H. F. Chau, Is quantum bit commitment really possible? Phys. Rev. Lett. **78**, 3410 (1997).

[24] D. Mayers, Unconditionally secure quantum bit commitment is impossible, Phys. Rev. Lett. **78**, 3414 (1997).

[25] U. Maurer and R. Renner, in *International Conference on Supercomputing* (Association for Computing Machinery (ACM), Tucson, AZ, USA, 2011).

[26] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner, Cryptography in the bounded-quantum-storage model, SIAM J. Comput. **37**, 1865 (2008).

[27] C. Erven, N. Ng, N. Gigov, R. Laflamme, S. Wehner, and G. Weihs, An experimental implementation of oblivious transfer in the noisy storage model, Nat. Commun. **5**, 3418 (2014).

[28] Y.-K. Liu, in *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science* (Association for Computing Machinery (ACM), Princeton, NJ, USA, 2014), pp. 269–286.

[29] F. Furrer, T. Gehring, C. Schaffner, C. Pacher, R. Schnabel, and S. Wehner, Continuous-variable protocol for oblivious transfer in the noisy-storage model, Nat. Commun. **9**, 1450 (2018).

[30] D. Mansy and P. Rindal, in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (Association for Computing Machinery (ACM), Princeton, NJ, USA, 2019), pp. 309–326.

[31] B. Mi, D. Huang, S. Wan, L. Mi, and J. Cao, Oblivious transfer based on ntruencrypt, IEEE Access **6**, 35283 (2018).

[32] B. Mi, D. Huang, S. Wan, Y. Hu, and K.-K. R. Choo, A post-quantum light weight 1-out-n oblivious transfer protocol, Comput. Electr. Eng. **75**, 90 (2019).

[33] P. Branco, L. Fiolhais, M. Goulão, P. Martins, P. Mateus, and L. Sousa, in *IACR Transactions on Cryptographic Hardware and Embedded Systems* (Association for Computing Machinery (ACM), London, UK, 2021), p. 215.

[34] S. Halevi and S. Micali, in *Advances in Cryptology–CRYPTO'96: 16th Annual International Cryptology Conference Santa Barbara, California, USA August 18–22, 1996 Proceedings 16* (Springer, Santa Barbara, CA, USA, 1996), pp. 201–215.

[35] M. Naor, Bit commitment using pseudorandomness, J. Cryptol. **4**, 151 (1991).

[36] R. Impagliazzo, in *Proceedings of Structure in Complexity Theory. Tenth Annual IEEE Conference* (IEEE, Minneapolis, MN, USA, 1995), pp. 134–147.

[37] B. Barak, in *Tutorials on the Foundations of Cryptography: Dedicated to Oded Goldreich* (Springer, New York, NY, USA, 2017), pp. 45–77.

[38] A. B. Grilo, H. Lin, F. Song, and V. Vaikuntanathan, in *Advances in Cryptology–EUROCRYPT 2021: 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17–21, 2021, Proceedings, Part II* (Springer, Zagreb, Croatia, 2021), pp. 531–561.

[39] A. C.-C. Yao, in *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery (ACM), Las Vegas, NV, USA, 1995), pp. 67–75.

[40] I. Damgård, S. Fehr, C. Lunemann, L. Salvail, and C. Schaffner, in *Annual International Cryptology Conference* (Springer, Santa Barbara, CA, USA, 2009), pp. 408–427.

[41] D. Unruh, in *EUROCRYPT* (Springer, Monaco and Nice, France, 2010), Vol. 6110, pp. 486–505.

[42] J. Bartusek, A. Coladangelo, D. Khurana, and F. Ma, in *Advances in Cryptology–CRYPTO 2021: 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16–20, 2021, Proceedings, Part I 41* (Springer, Santa Barbara, CA, USA, 2021), pp. 467–496.

[43] M. B. Santos, P. Mateus, and A. N. Pinto, Quantum oblivious transfer: A short review, Entropy **24**, 945 (2022).

[44] R. Konig, S. Wehner, and J. Wullschleger, Unconditional security from noisy quantum storage, IEEE Trans. Inf. Theory **58**, 1962 (2012).

[45] T. Kim, M. Fiorentino, and F. N. Wong, Phase-stable source of polarization-entangled photons using a polarization Sagnac interferometer, Phys. Rev. A **73**, 012316 (2006).

[46] M. Lemus, P. Schiansky, M. Goulão, M. Bozzio, D. Elkouss, N. Paunkovic, P. Mateus, and P. Walther, Experimental data accompanying publication "Performance of practical quantum oblivious key distribution", https://doi.org/10.5281/zenodo.16947845 (2025).

[47] T. Yamakawa and M. Zhandry, in *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (Springer, Monaco and Nice, France, 2021), pp. 568–597.

[48] C. Schaffner, Simple protocols for oblivious transfer and secure identification in the noisy-quantum-storage model, Phys. Rev. A: At. Mol. Opt. Phys. **82**, 032308 (2010).

[49] C. Greganti, P. Schiansky, I. A. Calafell, L. M. Procopio, L. A. Rozema, and P. Walther, Tuning single-photon sources for telecom multi-photon experiments, Opt. Express **26**, 3286 (2018).

[50] J. Hanel, Z. Jiang, J. Wang, F. Benthin, T. Fandrich, E. P. Rugeramigabo, R. Joos, M. Jetter, S. L. Portalupi, J. Yang, M. Zopf, P. Michler, and F. Ding, Ultrastable, low-error dynamic polarization encoding of deterministically generated single photons, arXiv:2507.16578 [quant-ph].

[51] R.-B. Jin, R. Shimizu, I. Morohashi, K. Wakui, M. Takeoka, S. Izumi, T. Sakamoto, M. Fujiwara, T. Yamashita, S. Miki, H. Terai, Z. Wang, and M. Sasaki, Efficient generation of twin photons at telecom wavelengths with 2.5 GHz repetition-rate-tunable comb laser, Sci. Rep. **4**, 7468 (2014).

[52] K. Wakui, Y. Tsujimoto, M. Fujiwara, I. Morohashi, T. Kishimoto, F. China, M. Yabuno, S. Miki, H. Terai, M. Sasaki, and M. Takeoka, Ultra-high-rate nonclassical light source with 50 GHz-repetition-rate mode-locked pump pulses and multiplexed single-photon detectors, Opt. Express **28**, 22399 (2020).

[53] M. Bozzio, A. Cavaillès, E. Diamanti, A. Kent, and D. Pitalúa-García, Multiphoton and side-channel attacks in mistrustful quantum cryptography, PRX Quantum **2**, 030338 (2021).

[54] A. M. Brańczyk, T. C. Ralph, W. Helwig, and C. Silberhorn, Optimized generation of heralded Fock states using parametric down-conversion, New J. Phys. **12**, 063001 (2010).

[55] B. Costa, P. Branco, M. Goulão, M. Lemus, and P. Mateus, Randomized oblivious transfer for secure multiparty computation in the quantum setting, Entropy **23**, 1001 (2021).

[56] O. Goldreich, *Foundations of Cryptography* (Cambridge University Press, USA, 2006), Vol. 1.

[57] D. Pitalúa-García, Spacetime-constrained oblivious transfer, Phys. Rev. A **93**, 062346 (2016).

[58] D. Unruh, in *International Conference on the Theory and Application of Cryptology and Information Security* (Springer, Hanoi, Vietnam, 2016), pp. 166–195.

[59] M. Zhandry, in *Annual International Cryptology Conference* (Springer, Santa Barbara, CA, USA, 2022), pp. 596–624.

[60] A. Agarwal, J. Bartusek, D. Khurana, and N. Kumar, in *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (Springer, Lyon, France, 2023), pp. 363–394.

[61] P. Mateus, A. Sernadas, and A. Souto, Universality of quantum Turing machines with deterministic control, J. Log. Comput. **27**, 1 (2017).

[62] C. W. Helstrom and C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976), vol. 84.

[63] C. A. Fuchs and J. Van De Graaf, Cryptographic distinguishability measures for quantum-mechanical states, IEEE Trans. Inf. Theory **45**, 1216 (1999).

[64] R. Renner, Security of quantum key distribution, Int. J. Quantum Inf. **6**, 1 (2008).

[65] W. Hoeffding, Probability inequalities for sums of bounded random variables, J. Am. Stat. Assoc. **58**, 13 (1963).

[66] N. J. Bouman and S. Fehr, in *Annual Cryptology Conference* (Springer, Santa Barbara, CA, USA, 2010), pp. 724–741.

[67] T. Lorünser, S. Ramacher, and F. Valbusa, Commitment schemes from OWFs with applications to quantum oblivious transfer, Entropy **27**, 751 (2025).

Macroscopic distinguishability between quantum states defining different phases of matter: Fidelity and the Uhlmann geometric phase

Nikola Paunković
SQIG - Instituto de Telecomunicações, IST, Lisbon, P-1049-001 Lisbon, Portugal

Vitor Rocha Vieira
CFIF and Department of Physics, IST, Technical University of Lisbon, P-1049-001 Lisbon, Portugal
(Received 31 July 2007; published 30 January 2008)

We study the fidelity approach to quantum phase transitions (QPTs) and apply it to general thermal phase transitions (PTs). We analyze two particular cases: The Stoner-Hubbard itinerant electron model of magnetism and the BCS theory of superconductivity. In both cases we show that the sudden drop of the mixed state fidelity marks the line of the phase transition. We conduct a detailed analysis of the general case of systems given by mutually commuting Hamiltonians, where the nonanalyticity of the fidelity is directly related to the nonanalyticity of the relevant response functions (susceptibility and heat capacity), for the case of symmetry-breaking transitions. Further, on the case of BCS theory of superconductivity, given by mutually noncommuting Hamiltonians, we analyze the structure of the system's eigenvectors in the vicinity of the line of the phase transition showing that their sudden change is quantified by the emergence of a generically nontrivial Uhlmann mixed state geometric phase.

DOI: 10.1103/PhysRevE.77.011129 PACS number(s): 05.70.Fh, 03.67.−a, 75.40.Cx

One of the main characteristics of quantum mechanics that makes it different from any classical physical theory is that in quantum mechanics two different quantum states, being either pure or mixed, are in general not fully distinguishable. By fully distinguishable we mean that it is possible, upon a result of a *single-shot* measurement of a suitable observable, to infer with probability one in which of the two given quantum states the observed system has been prepared. In particular, two pure quantum states are fully distinguishable if and only if they are orthogonal to each other. Otherwise, the maximal probability to unambiguously distinguish between two nonorthogonal pure quantum states is always strictly smaller than one. The reason for this lies in the fact that, while the outcomes of measurements on classical systems are, at least in principle, given with certainty, quantum measurements in general generate nontrivial probability distributions. This feature of quantum mechanics has found numerous applications within the field of quantum information and computation, in particular in quantum cryptography, quantum communication complexity, designing novel quantum algorithms, etc. (for an overview, see [1]).

Within the field of quantum information, the function widely used to quantify the distinguishability between two quantum states $\hat{\rho}_1$ and $\hat{\rho}_2$ is *fidelity* [2], given by the expression

$$F(\hat{\rho}_1,\hat{\rho}_2) = \mathrm{Tr}\sqrt{\sqrt{\hat{\rho}_1}\hat{\rho}_2\sqrt{\hat{\rho}_1}}. \tag{1}$$

Note that in the case of pure states $\hat{\rho}_1=|\psi_1\rangle\langle\psi_1|$ and $\hat{\rho}_2=|\psi_2\rangle\langle\psi_2|$, the above expression reduces to $F(|\psi_1\rangle\langle\psi_1|,|\psi_2\rangle\times\langle\psi_2|)=|\langle\psi_1|\psi_2\rangle|$, which is nothing but the square root of the probability for a system in state $|\psi_2\rangle$ to pass the test of being in state $|\psi_1\rangle$. The fidelity (1) between two quantum states, given for two systems 1 and 2, quantifies the statistical distinguishability between them, in a sense of *classical* statistical distinguishability between the probability distributions obtained by measuring an *optimal* observable in states $\hat{\rho}_1$ and $\hat{\rho}_2$. In other words, for every observable $\hat{A}$, we have $F(\hat{\rho}_1,\hat{\rho}_2)\leq F_c(\{p_1(i|\hat{A})\},\{p_2(i|\hat{A})\})\equiv\Sigma_i\sqrt{p_1(i|\hat{A})p_2(i|\hat{A})}$, where $\{p_\alpha(i|\hat{A})\}$, $\alpha\in\{1,2\}$, is a probability distribution obtained measuring the observable $\hat{A}$ in the state $\hat{\rho}_\alpha$, and $F_c(\{p_1(i|\hat{A})\},\{p_2(i|\hat{A})\})$ is the classical fidelity between the two probability distributions $\{p_1(i|\hat{A})\}$ and $\{p_2(i|\hat{A})\}$. For an overview of the results on distinguishability between quantum states and its applications to the field of quantum information, see [3] and the references therein.

Quantum mechanics was originally developed to describe the behavior of microscopic systems. Therefore, most of its applications are focused on the study of the properties and dynamics of quantum states referring to such systems, where quantum features dominate. On the other hand, it is a common assumption that classical behavior emerges in the thermodynamical limit, when the number of degrees of freedom of the system becomes large (in the limit of a large number of microscopic subsystems). Yet, there is no special objection why quantum mechanics should not be generally applicable, even to macroscopic systems. Indeed, macroscopic phenomena such as magnetism, superconductivity or superfluidity, to name just a few, can only be explained by using the rules of quantum mechanics. As in these, as well as in many other, highly physical relevant cases, the macroscopic features of matter are given through the features of its quantum states, the question of quantifying those macroscopic properties given by many-body quantum states arises as a relevant problem in physics.

In this study, we are interested in those macroscopic features of matter that define its thermodynamical phase. Different phases of matter are separated by the so-called *regions of criticality*, regions in parameter space where the system's

 ©2008 The American Physical Society

free energy becomes nonanalytic. As the free energy is a function of the system's states, it is precisely the features of its states that determine the phase the system is in. Indeed, different phases have different values of the *order parameter*, given by the expectation value of a certain observable. As in the case of general quantum states, here as well, fidelity can be used as a function whose behavior can mark the regions of criticality (and therefore the phase transitions).

In the case of quantum phase transitions (QPTs) [4], which occur at zero temperature and are driven by purely quantum fluctuations, the study of the ground state fidelity has been first conducted on the examples of the Dicke and *XY* models [5]. Note that in this case, the ground states are pure quantum states, so the fidelity is given by a simple overlap between two pure states. It was shown that approaching the regions of criticality the fidelity between two neighboring ground states exhibits a dramatic drop [26]. Subsequently, the fidelity approach to QPTs was applied to free Fermi systems and graphs [6] and to the Bose-Hubbard model [7]. The connection between fidelity, scaling behavior in QPTs and the renormalization group flows was introduced in [8] and further discussed in [9]. Also, it was shown that the fidelity can mark the regions of criticality in systems whose QPTs cannot be described in terms of Landau-Ginzburg-Wilson (LGW) theory: When the ground states are given by the matrix product states [10], in the case of a topologically ordered QPT [11], and in the Kosterlitz-Thouless type of transition [12] as well. The formal differential-geometry description of the fidelity approach to QPTs was first introduced in [13] and subsequently developed in [14], where the connection to the Berry phase approach to QPTs (see [15]) was also established. Further, on the example of the spin one-half *XXZ* Heisenberg chain, it was shown that the fidelity does not necessarily exhibit a dramatic drop at the critical point, but that the proper finite-size scaling analysis allows for correct identification of the QPT; see [14]. An interesting example of a Heisenberg chain where the fidelity approach fails when applied to the ground states, but does mark the point of criticality when applied to the first excited states, was discussed in [16] (note that the numerical results were obtained for up to 12 spins only, which leaves open the question of the ground state fidelity behavior in the thermodynamic limit). Introducing the temperature as an additional parameter, QPTs were studied in [17] and it was shown that extending the fidelity approach to general mixed (thermal) states can still mark the regions of criticality as well as the cross-over regions at finite temperatures. Finally, the genuine thermal PTs were discussed in [13] and [18], where the connection between the singularities in fidelity and specific heat or magnetic susceptibility was explicitly shown for the cases of systems given by mutually commuting Hamiltonians and symmetry-breaking PTs of LGW type.

In this paper, we apply the fidelity approach to general thermal phase transitions [27]. We analyze in detail two particular examples given by the Stoner-Hubbard model for magnetism and the BCS theory of superconductivity. In general, a system is defined by a Hamiltonian $\hat{H}(U)$ which is a function of a set of parameters representing the interaction coupling constants generically denoted as U. In thermal equilibrium, a system's state is given by a density operator $\hat{\rho}(T,U)$. Thus, in discussing the general, thermal as well as quantum phase transitions, we can consider the coupling constant(s) U and the temperature T to form a "generalized" parameter $q=(T,U)$. We consider the behavior of the fidelity $F(\hat{\rho}(q),\hat{\rho}(\tilde{q}))$ between two equilibrium thermal states $\hat{\rho}(q)$ and $\hat{\rho}(\tilde{q})$ defined by two close parameter points $q=q(T,U)$ and $\tilde{q}=q+\delta q=(T+\delta T,U+\delta U)$. We show that in both models considered, the PTs are marked by the sudden drop of fidelity in the vicinity of regions of criticality—a signature of enhanced distinguishability between two quantum states defining two different phases of matter, based on both short-range microscopic as well as long-range macroscopic features. For the general case of mutually commuting Hamiltonians, we analytically prove in detail that the same holds for PTs which fall within the symmetry-breaking paradigm described by the LGW theory (see also [13] and [18]). Further, for the case of mutually noncommuting Hamiltonians, on the example of BCS theory of superconductivity we show that the nonanalyticity of the fidelity is accompanied by the emergence of a generically nontrivial Uhlmann geometric phase [19], the mixed-state generalization of the Berry geometric phase (for the relation between QPTs and Berry phases, see [15], [13], and [14]).

I. STONER-HUBBARD ITINERANT ELECTRON MODEL FOR MAGNETISM

First, we discuss the case of the Stoner-Hubbard model for itinerant electrons on a lattice given by the Hamiltonian [20]:

$$\hat{H}_{\mathrm{SH}}=\sum_{k}\varepsilon_{k}(\hat{c}_{k\uparrow}^{\dagger}\hat{c}_{k\uparrow}+\hat{c}_{k\downarrow}^{\dagger}\hat{c}_{k\downarrow})+U\sum_{l}\hat{c}_{l\uparrow}^{\dagger}\hat{c}_{l\uparrow}\hat{c}_{l\downarrow}^{\dagger}\hat{c}_{l\downarrow}. \quad (2)$$

The anticommuting fermionic operators $\hat{c}_{k\sigma}^{\dagger}$ represent the free-electron momentum Bloch modes ($\sigma\in\{\uparrow,\downarrow\}$), while the on-site operators $\hat{c}_{l\sigma}^{\dagger}=V^{-1/2}\Sigma_{k}e^{-ikx_{l}}\hat{c}_{k\sigma}^{\dagger}$ are given by their Fourier transforms, where x_l represents the position of the lth lattice site. The coupling constant $U>0$ defines the on-site electron Coulomb repulsion. There are in total N electrons and they occupy the volume V (such that in the thermodynamic limit, when $N,V\to\infty$, we have $N/V\to\mathrm{const}$). Finally, we assume for simplicity that the kinetic energy is given by $\varepsilon_{k}=\hbar^{2}k^{2}/(2m)$, while the number operators are $\hat{n}_{k\sigma}=\hat{c}_{k\sigma}^{\dagger}\hat{c}_{k\sigma}$ and $\hat{n}_{l\sigma}=\hat{c}_{l\sigma}^{\dagger}\hat{c}_{l\sigma}=V^{-1}\Sigma_{qk}e^{iqx_{l}}\hat{c}_{k\sigma}^{\dagger}\hat{c}_{k+q\sigma}$. Thus

$$\hat{H}_{\mathrm{SH}}=\sum_{k\sigma}\varepsilon_{k}\hat{n}_{k\sigma}+U\sum_{l}\hat{n}_{l\uparrow}\hat{n}_{l\downarrow}. \quad (3)$$

In order to obtain the mean-field effective Hamiltonian, we neglect the term quadratic in the fluctuations $\delta\hat{n}_{l\sigma}=\hat{n}_{l\sigma}-n_{l\sigma}$, where $n_{l\sigma}=\langle\hat{n}_{l\sigma}\rangle$, in the potential written as $\hat{n}_{l\uparrow}\hat{n}_{l\downarrow}=\delta\hat{n}_{l\uparrow}\delta\hat{n}_{l\downarrow}+n_{l\uparrow}\hat{n}_{l\downarrow}+n_{l\downarrow}\hat{n}_{l\uparrow}-n_{l\uparrow}n_{l\downarrow}$. Expressing the number operators in terms of the Bloch momentum operators, and using

$\langle \hat{c}^{\dagger}_{k\sigma}\hat{c}_{k'\sigma}\rangle = \delta_{kk'}\langle \hat{n}_{k\sigma}\rangle$ (which expresses the translational invariance in a ferromagnetic ground state), the mean-field linearized effective Hamiltonian becomes

$$\hat{H}^{\text{eff}}_{\text{SH}} = \sum_k (E_{k\uparrow}\hat{n}_{k\uparrow} + E_{k\downarrow}\hat{n}_{k\downarrow}) - VUn_{\uparrow}n_{\downarrow}. \tag{4}$$

Here, $n_\sigma = N_\sigma / V$ is the density of electrons with spin projection along z axis given by $\sigma \in \{\uparrow, \downarrow\}$ ($N_\sigma = \Sigma_k \langle \hat{n}_{k\sigma}\rangle$ being the total number of electrons with spin σ). The one-particle electron energies in this effective model are obtained by shifting the free-electron energies ε_k by an amount depending on the particle's spin:

$$E_{k\uparrow} = \varepsilon_k + Un_{\downarrow},$$

$$E_{k\downarrow} = \varepsilon_k + Un_{\uparrow}. \tag{5}$$

Since the one-particle energy modes are decoupled, the overall ground state is obtained by filling the electrons up to the Fermi level ε_F:

$$|g\rangle = \otimes_{k\le k_{F\uparrow}} \hat{c}^{\dagger}_{k\uparrow} \otimes_{k\le k_{F\downarrow}} \hat{c}^{\dagger}_{k\downarrow} |0\rangle, \tag{6}$$

where $|0\rangle$ represents the vacuum state with no electrons and $k_{F\uparrow}$ the maximal value of the momentum for spin up electrons, given by $E_{k_{F\uparrow}} = \varepsilon_F$ (and analogously for $k_{F\downarrow}$) [28]. Note that, due to the different dispersion formulas (5) for particles with spin up and spin down, in general the values of $k_{F\uparrow}$ and $k_{F\downarrow}$ that minimize the ground state energy, and therefore define the state (6) itself, are different and consequently the number of up and down electrons will be different. This is precisely the reason for the existence of magnetism in this model. As soon as the energy of the "biased" (magnetic) state, for which for example $k_{F\uparrow} > k_{F\downarrow}$, becomes lower than the energy of the "balanced" (paramagnetic) state ($k_{F\uparrow} = k_{F\downarrow}$), a magnetic phase transition will occur. Obviously, for reasons of symmetry, the magnetic state can be reversed with the $k_{F\uparrow} > k_{F\downarrow}$ and $k_{F\uparrow} < k_{F\downarrow}$ cases having the same energy, in the absence of an external symmetry breaking field $\vec{H}$. The qualitative picture of the emergence of magnetic features can be seen already from looking at the original Hamiltonian (2): In the $U \to 0$ limit, when the Coulomb interaction is negligible, the Hamiltonian represents a system of free electrons that exhibits no magnetic order (all Bloch states are doubly occupied); in the opposite $U \to \infty$ limit, the second term of the Hamiltonian becomes the dominant one and is minimized by one of two possible states for which either $N_\uparrow = 0$ or $N_\downarrow = 0$.

The quantitative analysis of the zero-temperature critical behavior of the effective Hamiltonian (4) can be done by looking at the divergence of the magnetic susceptibility, using the one-particle energy dispersion relations (5). At $T=0$ it leads to the well known Stoner criterion for the emergence of magnetism [20]:

$$D_F U_c = 1, \tag{7}$$

where U_c is the critical value of the coupling constant above which the system is in a magnetic phase and $D_F = D(\varepsilon_F)$ is the density of states around the Fermi energy. From this, one can obtain (see Appendix A) the critical value U_c of the coupling constant [29]:

$$U_c = \frac{4}{3}\varepsilon_F \frac{V}{N}, \tag{8}$$

where N is the total number of electrons, V is the volume of the system, and m is the electron mass.

Alternatively, the above result can be derived by minimizing the overall ground state energy, thus obtaining the explicit dependencies $k_{F\uparrow} = k_{F\uparrow}(U)$ and $k_{F\downarrow} = k_{F\downarrow}(U)$ which determine the ground state (6) and U_c in particular (the maximum value of U for which $k_{F\uparrow} = k_{F\downarrow}$) [30]. The total number of electrons in the system is given by $N = N_\uparrow + N_\downarrow$, and thus

$$\frac{N}{V} = \frac{1}{6\pi^2}[(k_{F\uparrow})^3 + (k_{F\downarrow})^3], \tag{9}$$

since the up and down electrons occupy spheres of radius $k_{F\uparrow}$ and $k_{F\downarrow}$, with volumes $V(k_{F\uparrow})$ and $V(k_{F\downarrow})$, in momentum space, with a density of states $V/(2\pi)^3$, as follows from the periodic boundary conditions for the Bloch functions. From $E_{k_{F\uparrow}} = E_{k_{F\downarrow}}$ (see Appendix A), using the energy dispersion formulas (5), we obtain the second equation that determines the Fermi momenta $k_{F\uparrow}$ and $k_{F\downarrow}$ (with $\alpha = \frac{3}{2}k_F U_c$):

$$(k_{F\uparrow} - k_{F\downarrow})\left[(k_{F\uparrow} + k_{F\downarrow}) - \frac{U}{\alpha}(k^2_{F\uparrow} + k_{F\uparrow}k_{F\downarrow} + k^2_{F\downarrow})\right] = 0. \tag{10}$$

We see that the above two equations always have the trivial solution $k_{F\uparrow} = k_{F\downarrow} \equiv k_F$, which from Eq. (9) is $k_F = [3\pi^2(N/V)]^{1/3}$. Yet, it is not necessarily the only possible solution, as the quadratic term in Eq. (10) can also be satisfied: $(k_{F\uparrow} + k_{F\downarrow}) - (U/\alpha)(k^2_{F\uparrow} + k_{F\uparrow}k_{F\downarrow} + k^2_{F\downarrow}) = 0$. It turns out that precisely for $U > U_c$ this term has nontrivial, "nonbalanced" solutions that are energetically more favorable than the "balanced" one and that give rise to magnetic order (see Appendix A).

Thus Eqs. (9) and (10) define $k_{F\uparrow} = k_{F\uparrow}(U)$ and $k_{F\downarrow} = k_{F\downarrow}(U)$ functions which, via Eq. (6), give the ground state as a function of the external parameter U, $|g\rangle = |g(U)\rangle$. This enables us to analyze the fidelity between two ground states $|g\rangle \equiv |g(U)\rangle$ and $|\tilde{g}\rangle = |g(U+\delta U)\rangle$ in two close parameter points U and $U+\delta U$. The fidelity is then $F(|g\rangle\langle g|, |\tilde{g}\rangle\langle\tilde{g}|) = |\langle g|\tilde{g}\rangle|$, and for $U < U_c$ (and δU sufficiently small, i.e., $\delta U < U_c - U$) we see that the fidelity is identical to 1. This is a simple consequence of our mean-field approximation based on a simplified description in terms of single particle energy states. On the magnetic side of the phase diagram, the ground states are indeed different from each other, which follows from the relation $k_{F\uparrow}(U) \neq k_{F\downarrow}(U)$. In fact, from Eq. (6) it follows that *any* two ground states with different numbers $N_\uparrow$ and $N_\downarrow$ are *orthogonal* to each other. Since this is precisely the case in the thermodynamic limit, the fidelity between any two different ground states (in two different parameter points) is identically equal to zero. This is the

famous Anderson orthogonality catastrophe, discussed in more detail in [5]. For systems with infinitely many degrees of freedom, such as those taken in the thermodynamic limit, every two ground states are generally orthogonal to each other. Therefore, in order to infer the points of criticality, we are forced to either analyze the finite-size scaling behavior (see [5]), or to introduce the fidelity per lattice site and work directly in the thermodynamic limit (see [8], [9], and [21]).

In our case though, even for finite systems the ground state is discontinuous at the point of criticality. This is due to the fact that the unperturbed and the symmetry-breaking perturbation Hamiltonian commute with each other, which results in a first-order quantum phase transition at the point of level-crossing between the ground and the first excited state. In other words, in the case of finite systems $(N, V<\infty)$, the small enough changes of the parameter $U \to U+\delta U$ will result in small changes of volumes $V(k_{F\uparrow}) \to V(k_{F\uparrow})+\delta V(k_{F\uparrow})$ such that $|\delta V(k_{F\uparrow})| < V_0=(2\pi)^3/V$ (V_0 is the volume occupied by each one-particle state in momentum space): Infinitesimal changes of the volumes $V(k_{F\uparrow})$ and $V(k_{F\downarrow})$ are small enough to cause a change in the numbers $N_\uparrow$ and $N_\downarrow$, and thus in the ground state. Therefore, for finite systems, the fidelity between two ground states is either one or zero—it is not a continuous function and its rate of change cannot be analyzed directly.

Yet, we can use the rate of change of $V(k_{F\uparrow})$, the derivative $dV(k_{F\uparrow})/dU$, to quantify the change of fidelity itself [note that, due to the fixed total number of electrons N, we have that $dV(k_{F\downarrow})/dU=-dV(k_{F\uparrow})/dU$]. Lengthy, but elementary algebra (see Appendix A) shows that precisely at $U=U_c$, the derivative $\frac{dV(k_{F\uparrow})}{dU}$ diverges to infinity, thus marking the *macroscopic* distinguishability between the states from paramagnetic and magnetic phase (see Fig. 1).

Next, we discuss the general case of $T\neq 0$ phase transitions. First, we transform the Hamiltonian $\hat{H}^{\text{eff}}_{\text{SH}}$ to a form with the explicit symmetry breaking term that drives the phase transition. Using $M=(N_\uparrow-N_\downarrow)/2=\langle\hat{S}^z\rangle=\Sigma_k\langle\hat{S}^z_k\rangle$, with $N_\sigma=\langle\hat{N}_\sigma\rangle=\Sigma_k\langle\hat{n}_{k\sigma}\rangle$, $\hat{n}_k=\hat{n}_{k\uparrow}+\hat{n}_{k\downarrow}$, and $\hat{S}^z_k=\frac{1}{2}(\hat{n}_{k\uparrow}-\hat{n}_{k\downarrow})=\hat{\Psi}^\dagger(\sigma^z/2)\hat{\Psi}$, where $\hat{\Psi}^\dagger_k=[\hat{c}^\dagger_{k\uparrow}\hat{c}^\dagger_{k\downarrow}]$ and σ^z is the z component of the vector $\vec{\sigma}$ of Pauli matrices, we get

$$\hat{H}^{\text{eff}}_{\text{SH}}=\sum_k\left[\left(\varepsilon_k+\frac{UN}{2V}\right)\hat{n}_k-2\frac{UM}{V}\hat{S}^z_k\right]-\frac{U}{V}\left(\frac{N^2}{4}-M^2\right). \tag{11}$$

In thermal equilibrium, the state of the system is given by $\hat{\rho}=(1/Z)e^{-\beta(\hat{H}^{\text{eff}}_{\text{SH}}-\mu\hat{N})}$, where $Z=\text{Tr}[e^{-\beta(\hat{H}^{\text{eff}}_{\text{SH}}-\mu\hat{N})}]$ is the grand canonical partition function, $\beta=1/(k_BT)$, with k_B being the Boltzmann constant and T the absolute temperature and $\mu=\mu(T)$ is the chemical potential. Using the above expression for the Hamiltonian $\hat{H}^{\text{eff}}_{\text{SH}}$, we can write

$$-\beta(\hat{H}^{\text{eff}}_{\text{SH}}-\mu\hat{N})=\sum_k(\alpha_k\hat{n}_k+h_z\hat{S}^z_k)+C. \tag{12}$$

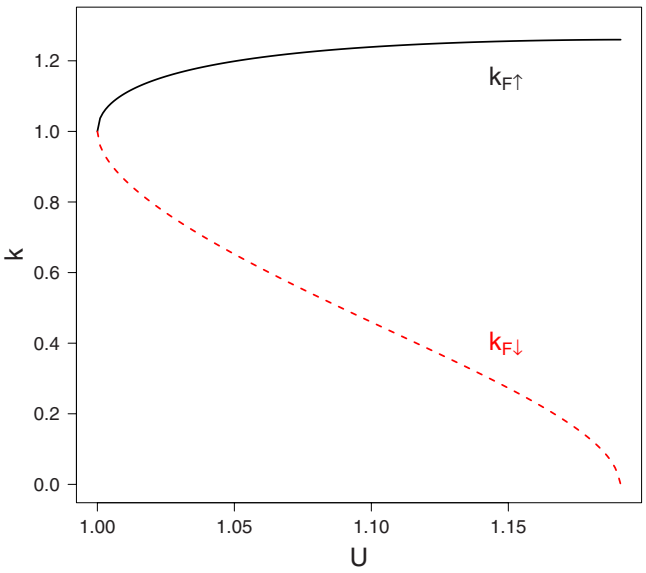


FIG. 1. (Color online) Explicit dependencies $k_{F\uparrow}=k_{F\uparrow}(U)$ (solid line) and $k_{F\downarrow}=k_{F\downarrow}(U)$ (dashed line) determining the ground state (6). Note that at the point of QPT, given by $U_c=1$, the two derivatives $dk_{F\uparrow}/dU$ and $dk_{F\downarrow}/dU$ become infinite, marking the nonanalyticity in the ground state fidelity. The plot is given in rescaled quantities $U \to D_F U$ and $k_{F\sigma} \to k_{F\sigma}/k_F$.

Here, $\alpha_k=-\beta E_k$ (with $E_k=\bar{\varepsilon}_k+UN/2V$ and $\bar{\varepsilon}_k=\varepsilon_k-\mu$), $h_z=2\beta(U/V)M$, and $C=\beta(U/V)(N^2/4-M)$. Note that the coefficients α_k, h_z, and C are functions of both the coupling constant U and, through β and the chemical potential $\mu=\mu(T)$, of the temperature T as well, so that the "generalized" parameter is $q=(T,U)$. Using the obvious commutation relations $[\hat{n}_k,\hat{n}_{k'}]=[\hat{n}_k,\hat{S}^z_{k'}]=[\hat{S}^z_k,\hat{S}^z_{k'}]=0$ (for $k\neq k'$), the equilibrium state can be expressed as

$$\hat{\rho}=\frac{1}{Z}e^{-\beta(\hat{H}^{\text{eff}}_{\text{SH}}-\mu\hat{N})}=\frac{e^{\sum_k(\alpha_k\hat{n}_k+h_z\hat{S}^z_k)+C}}{\text{Tr}\left[e^{\sum_k(\alpha_k\hat{n}_k+h_z\hat{S}^z_k)+C}\right]}=\frac{\prod_k(e^{\alpha_k\hat{n}_k}e^{h_z\hat{S}^z_k})}{\prod_k\text{Tr}[e^{\alpha_k\hat{n}_k}e^{h_z\hat{S}^z_k}]}. \tag{13}$$

We next choose two parameter points $q_a=(T_a,U_a)$ and $q_b=(T_b,U_b)$ defining the Hamiltonians $\hat{H}_a=\hat{H}^{\text{eff}}_{\text{SH}}(U_a)$ and $\hat{H}_b=\hat{H}^{\text{eff}}_{\text{SH}}(U_b)$ and the corresponding equilibrium states $\hat{\rho}_a=\hat{\rho}(q_a)$ and $\hat{\rho}_b=\hat{\rho}(q_b)$, respectively. The fidelity between the two states is then given by

$$F(\hat{\rho}_a,\hat{\rho}_b)=\text{Tr}[(\hat{\rho}_a^{1/2}\hat{\rho}_b\hat{\rho}_a^{1/2})^{1/2}]=\text{Tr}[\sqrt{\hat{\rho}_a\hat{\rho}_b}]$$
$$=\frac{\text{Tr}[e^{(\beta_a\hat{H}_a+\beta_b\hat{H}_b)/2}]}{\sqrt{Z(\hat{H}_a)Z(\hat{H}_b)}}. \tag{14}$$

Using Eq. (13) and the expression $\text{Tr}[e^{(\alpha_k\hat{n}_k+h_z\hat{S}^z_k)}]=2e^{\alpha_k}[\cosh\alpha_k+\cosh(h_z/2)]$ (for the proof, see Appendix B), the fidelity between two different equilibrium states $\hat{\rho}_a$ and $\hat{\rho}_b$ finally becomes

$$F(\hat{\rho}_a,\hat{\rho}_b)=\prod_k \frac{\cosh\bar{\alpha}_k+\cosh\frac{\bar{h}_z}{2}}{\sqrt{\left[\cosh\left(\bar{\alpha}_k+\frac{\Delta\alpha_k}{2}\right)+\cosh\left(\frac{\bar{h}_z}{2}+\frac{\Delta h_z}{2}\right)\right]\left[\cosh\left(\bar{\alpha}_k-\frac{\Delta\alpha_k}{2}\right)+\cosh\left(\frac{\bar{h}_z}{2}-\frac{\Delta h_z}{2}\right)\right]}}, \tag{15}$$

with $\bar{\alpha}_k=[\alpha_k(q_a)+\alpha_k(q_b)]/2$, $\Delta\alpha_k=\alpha_k(q_a)-\alpha_k(q_b)$, and similarly for $\bar{h}_z$ and Δh_z. If we choose the two points to be close to each other, $\Delta\alpha_k\ll 1$ and $\Delta h_z\ll 1$, then the fidelity can be seen as a function of $\bar{\alpha}_k$ and $\bar{h}_z$, with a fixed parameter difference.

In order to evaluate the fidelity (15), we need to determine the magnetization $M=M(T,U)$ and the chemical potential $\mu=\mu(T,U)$, given by the pair of self-consistent integral equations:

$$N=VD_F\int_0^{+\infty}d\varepsilon\sqrt{\frac{\varepsilon}{\varepsilon_F}}[f(E_{k\uparrow})+f(E_{k\downarrow})],$$

$$M=VD_F\int_0^{+\infty}d\varepsilon\sqrt{\frac{\varepsilon}{\varepsilon_F}}\frac{1}{2}[f(E_{k\uparrow})-f(E_{k\downarrow})], \tag{16}$$

where $f(E_{k\sigma})=[\exp(\beta E_{k\sigma})+1]^{-1}$ is the usual Fermi distribution. We used the subroutine hybrd.f from MINPACK [22] to solve the above system numerically. In the $T\to 0$ limit, the above system reduces to Eqs. (9) and (10) that determine the $T=0$ ground state. The result for the magnetization is given in Fig. 2. The line of the phase transition $U_c=U_c(T)$ is clearly marked, and is plotted in Fig. 3. Finally, using the numerical results for M and μ, we obtain the fidelity, depicted in Fig. 4. We clearly see the same line of the phase transition as the line of a sudden drop of F. Note that all the plots are given in rescaled quantities $T\to k_BT$, $U\to D_FU$, and $M\to M/N$, with $\delta T=0$ and $\delta U=2\times 10^{-3}$. We have evaluated the fidelity for $\delta T=2\times 10^{-3}$ and $\delta U=0$, as well as for $\delta T=\delta U=2\times 10^{-3}$ and the results are qualitatively the same.

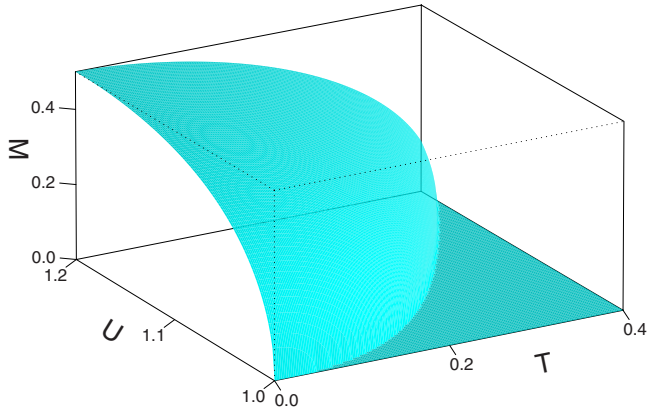


FIG. 2. (Color online) Magnetization $M=M(T,U)$ as a function of the temperature T and the coupling constant U. The plot is given in rescaled quantities $T\to k_BT$, $U\to D_FU$, and $M\to M/N$.

II. GENERAL CASE OF MUTUALLY COMMUTING HAMILTONIANS

The above singular behavior of the fidelity indeed marks the regions of phase transitions, not only for the case of the Stoner-Hubbard model but also for the broad class of systems given by a set of mutually commuting Hamiltonians $\hat{H}(q)$ and whose critical behavior is explained by the LGW theory, as can be seen by the following analysis. Using the mean Hamiltonian $\hat{\bar{H}}=(\hat{H}_a+\hat{H}_b)/2$ and the difference $\Delta\hat{H}=\hat{H}_a-\hat{H}_b$, from Eq. (14) the fidelity can be written as

$$F(\hat{\rho}_a,\hat{\rho}_b)=\frac{Z(\hat{\bar{H}})}{\sqrt{Z\left(\hat{\bar{H}}+\frac{\Delta\hat{H}}{2}\right)Z\left(\hat{\bar{H}}-\frac{\Delta\hat{H}}{2}\right)}}. \tag{17}$$

Within the LGW theory, PTs occur as a consequence of the emergence of a symmetry-breaking term in the Hamiltonian, given by an operator $\hat{S}$. Thus we can write the overall Hamiltonian as a sum of unperturbed and symmetry-breaking

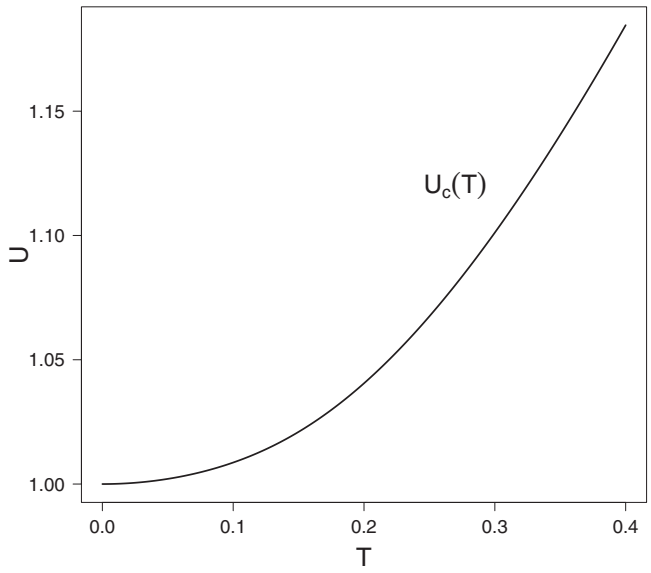


FIG. 3. Critical line of the magnetic phase transition $U_c=U_c(T)$. The plot is given in rescaled quantities $T\to k_BT$ and $U\to D_FU$.

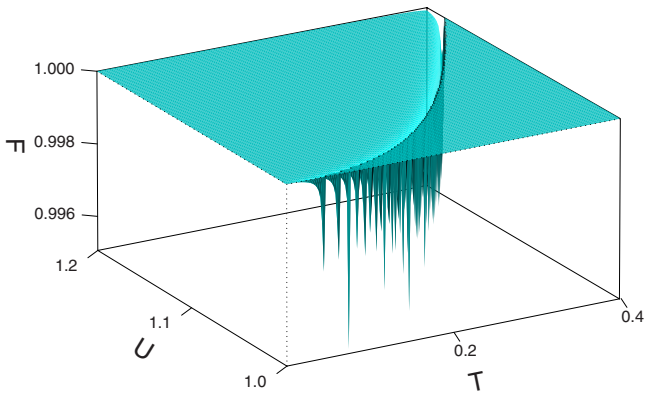


FIG. 4. (Color online) Fidelity $F=F(T,U)$. The plot is given in rescaled quantities $T\rightarrow k_BT$ and $U\rightarrow D_FU$, with $\delta T=0$ and $\delta U=2\times10^{-3}$.

terms, $\hat{H}=\hat{H}_0-h\hat{S}$, with $h=h(q)$. Using [31] $Z(\hat{H})=\mathrm{Tr}[e^{-\beta(\hat{H}_0-h\hat{S})}]$ (note the implicit dependence of partition function Z on temperature T, through β), for the first and the second derivative we obtain $\partial\ln Z/\partial h=\beta\,\mathrm{Tr}[\hat{\rho}(\hat{H})\hat{S}]=\beta\langle\hat{S}\rangle=\beta M$ and $(\partial^2/\partial h^2)(\ln Z)=\beta\chi$. Here, $M=M(h)$ and $\chi=\chi(h)=\beta(\langle\hat{S}^2\rangle-\langle\hat{S}\rangle^2)$ are the generalized "magnetization" and "susceptibility," respectively. Thus, for the logarithm of the partition function (free energy), we obtain $\ln Z(\hat{H})|_{h+\Delta h}=\ln Z(\hat{H})|_h+\beta M(h)\Delta h+\frac{1}{2}\beta\chi(h)\Delta h^2+o(\Delta h^2)$. Finally, in the vicinity of the points of parameter space where the self-consistent field h vanishes, the fidelity (17) reads as

$$F|_{h=\Delta h}\simeq e^{-(1/2)\beta\chi(0)\Delta h^2}. \tag{18}$$

According to the LGW theory, at the phase transition, the zeroth-field "susceptibility" $\chi(0)$ becomes nonanalytic and diverges. Thus it follows that the fidelity F will itself become nonanalytic and experience a sudden drop.

When interested in the system's behavior at PTs, we usually simplify the problem by taking the unperturbed Hamiltonian to be constant. Yet, in general, the unperturbed Hamiltonian is also a function of the parameters, $\hat{H}_0=\hat{H}_0(q)$. This gives the correction to the above formula for the fidelity which, introducing $\hat{\bar{H}}=\hat{\bar{H}}_0-\bar{h}\hat{S}$ and $\Delta\hat{H}=\Delta\hat{H}_0-\Delta h\hat{S}$, reads as

$$F|_{h=\Delta h}\simeq\frac{Z(\hat{\bar{H}}_0)}{\left[Z\left(\hat{\bar{H}}_0+\frac{\Delta\hat{H}_0}{2}\right)Z\left(\hat{\bar{H}}_0-\frac{\Delta\hat{H}_0}{2}\right)\right]^{1/2}}e^{-(1/2)\beta\chi(0)\Delta h^2}. \tag{19}$$

The "correction" term in the above product is responsible for "short-range" local correlations, while the second one quantifies the global "long-range" correlations giving rise to macroscopic phase distinguishability. Note though that even within a single phase, a system can be in different macroscopically distinguishable states—phase distinguishability is not the only form of macroscopic distinguishability. Yet, it is in some sense the "extreme" version of it, which clearly affects the behavior of fidelity.

Finally, we note that in the above discussion we focused on PTs driven by the local order parameter $\hat{S}$. Therefore, we analyzed the Taylor expansion of F with respect to Δh deviations only, which at the second order are given by the generalized susceptibility $\chi(h)$. In the general case, considering the temperature deviations as well, one would include additional terms involving the specific heat $C=C(q)$, again resulting in a singular behavior of the fidelity.

III. BCS SUPERCONDUCTIVITY

Next, we discuss the BCS theory for superconductivity [20], providing us with an example of a model with mutually noncommuting Hamiltonians. The one-electron Bloch momentum modes are given by the fermionic anticommuting operators $\hat{c}_{k\sigma}$ (label $\sigma\in\{\uparrow,\downarrow\}$ represents spins with projections up and down along, say z axis), with the one-particle kinetic energies taken to be, again for simplicity, $\varepsilon_k=\hbar^2k^2/(2m)$. The BCS superconducting Hamiltonian that represents the sum of one-particle kinetic and Cooper-pair interaction energies can be written in the following way ($V_{k'k}=V^*_{kk'}$ are the coupling constants):

$$\hat{H}_{\mathrm{BCS}}=\sum_{k\sigma}\varepsilon_k\hat{c}^\dagger_{k\sigma}\hat{c}_{k\sigma}+\sum_{kk'}V_{kk'}\hat{c}^\dagger_{k'\uparrow}\hat{c}^\dagger_{-k'\downarrow}\hat{c}_{-k\downarrow}\hat{c}_{k,\uparrow}. \tag{20}$$

By $\hat{n}_{k\sigma}=\hat{c}^\dagger_{k\sigma}\hat{c}_{k\sigma}$ we denote the one-particle number operators, while by $\hat{b}^\dagger_k=\hat{c}^\dagger_{k\uparrow}\hat{c}^\dagger_{-k\downarrow}$ and $\hat{b}_k=\hat{c}_{-k\downarrow}\hat{c}_{k\uparrow}$ we define the Cooper-pair creation and annihilation operators, respectively. Analogously to the previous case, using $\hat{b}_k=\langle\hat{b}_k\rangle+\delta\hat{b}_k$ and neglecting the term quadratic in the fluctuations, we obtain the effective mean-field BCS Hamiltonian:

$$\hat{H}^{\mathrm{eff}}_{\mathrm{BCS}}=\sum_k\varepsilon_k(\hat{n}_{k\uparrow}+\hat{n}_{-k\downarrow})-\sum_k(\Delta_k\hat{b}^\dagger_k+\Delta^*_k\hat{b}_k-\Delta^*_kb_k), \tag{21}$$

with $\Delta_k=-\Sigma_{k'}V_{kk'}b_{k'}$ and $b_k=\langle\hat{b}_k\rangle$. We will use the usual assumption that the lattice-mediated pairing interaction is constant and nonvanishing between electrons around the Fermi level only, i.e., $V_{kk'}=-V$ for $|\bar{\varepsilon}_k|$ and $|\bar{\varepsilon}_{k'}|<\hbar\omega_D$, and zero otherwise ($\omega_D$ is the Debye frequency). Using the Nambu operators [20] $\hat{\bar{T}}_k=\hat{\psi}^\dagger_k(\vec{\sigma}/2)\hat{\psi}_k$, where $\hat{\psi}^\dagger_k=[\hat{c}^\dagger_{k\uparrow}\hat{c}_{-k\downarrow}]$ and $\vec{\sigma}$ is the vector of Pauli matrices [32], the operators are given by $\hat{T}^+_k=\hat{b}^\dagger_k$, $\hat{T}^-_k=\hat{b}_k$ and $2\hat{T}^0_k+1=(\hat{n}_{k\uparrow}+\hat{n}_{-k\downarrow})$ and form a su(2) algebra (see Appendix C). Using this notation, the Hamiltonian takes the form

$$\hat{H}^{\mathrm{eff}}_{\mathrm{BCS}}=\sum_k(2\varepsilon_k\hat{T}^0_k-\Delta_k\hat{T}^+_k-\Delta^*_k\hat{T}^-_k)+\sum_k(\varepsilon_k+\Delta^*_kb_k). \tag{22}$$

As before, the thermal equilibrium state is given by $\hat{\rho}=(1/Z)e^{-\beta(\hat{H}^{\mathrm{eff}}_{\mathrm{BCS}}-\mu\hat{N})}$ and, using the above expression for the Hamiltonian $\hat{H}^{\mathrm{eff}}_{\mathrm{BCS}}$, we can write

$$-\beta(\hat{H}_{\mathrm{BCS}}^{\mathrm{eff}}-\mu\hat{N})=\sum_k \vec{\tilde{h}}_k\hat{\vec{T}}_k+K, \tag{23}$$

where $\vec{\tilde{h}}_k=(\tilde{h}_k^+,\tilde{h}_k^-,\tilde{h}_k^0)=(2\beta\Delta_k^*,2\beta\Delta_k,-2\beta\bar{\varepsilon}_k)$, $\hat{\vec{T}}_k=(\hat{T}_k^+,\hat{T}_k^-,\hat{T}_k^0)$, $K=-\beta\Sigma_k(\bar{\varepsilon}_k+\Delta_k^* b_k)$, and $\bar{\varepsilon}_k=\varepsilon_k-\mu(T)$. The norms of the vectors $\vec{\tilde{h}}_k$ are given by $\tilde{h}_k=2\beta E_k$, with $E_k=\sqrt{\bar{\varepsilon}_k^2+|\Delta_k|^2}$. Similarly to what we had before, the coefficients $\vec{\tilde{h}}_k=\vec{\tilde{h}}_k(T,V)$ are functions of both the coupling constant V and the temperature T, through the gap parameters $\Delta_k=\Delta_k(T,V)$ and the chemical potential $\mu=\mu(T)$. Thus we can talk of the "generalized" parameter $q=(T,V)$. Since $[\hat{\vec{T}}_k,\hat{\vec{T}}_{k'}]=0$, for $k\neq k'$, we have

$$\hat{\rho}=\frac{1}{Z}e^{-\beta(\hat{H}_{\mathrm{BCS}}^{\mathrm{eff}}-\mu\hat{N})}=\frac{e^{\Sigma_k\vec{\tilde{h}}_k\hat{\vec{T}}_k+K}}{\mathrm{Tr}[e^{\Sigma_k\vec{\tilde{h}}_k\hat{\vec{T}}_k+K}]}=\frac{\prod_k e^{\vec{\tilde{h}}_k\hat{\vec{T}}_k}}{\prod_k \mathrm{Tr}[e^{\vec{\tilde{h}}_k\hat{\vec{T}}_k}]}. \tag{24}$$

We wish to evaluate the fidelity between two thermal states $\hat{\rho}_a$ and $\hat{\rho}_b$, given for two different parameter points $q_a=(T_a,V_a)$ and $q_b=(T_b,V_b)$. Using definition (1) and $\vec{a}_k=\vec{\tilde{h}}_k(q_a)$ and $\vec{b}_k=\vec{\tilde{h}}_k(q_b)$, we have

$$F(\hat{\rho}_a,\hat{\rho}_b)=\mathrm{Tr}[(\hat{\rho}_a^{1/2}\hat{\rho}_b\hat{\rho}_a^{1/2})^{1/2}]=\frac{\mathrm{Tr}\left[\left(\prod_k e^{(\vec{a}_k/2)\hat{\vec{T}}_k}e^{\vec{b}_k\hat{\vec{T}}_k}e^{(\vec{a}_k/2)\hat{\vec{T}}_k}\right)^{1/2}\right]}{\prod_k(\mathrm{Tr}[e^{\vec{a}_k\hat{\vec{T}}_k}]\mathrm{Tr}[e^{\vec{b}_k\hat{\vec{T}}_k}])^{1/2}}. \tag{25}$$

As for every k the operators $\hat{\vec{T}}_k$ form a su(2) algebra, and therefore by exponentiation define a Lie group, we can write $e^{(\vec{a}_k/2)\hat{\vec{T}}_k}e^{\vec{b}_k\hat{\vec{T}}_k}e^{(\vec{a}_k/2)\hat{\vec{T}}_k}=e^{2\vec{c}_k\hat{\vec{T}}_k}$. Also (see Appendix C), we have that $\mathrm{Tr}[e^{\vec{a}_k\hat{\vec{T}}_k}]=2[1+\cosh(a_k/2)]$. Therefore, we finally have (see Appendix C) that $F(\hat{\rho}_a,\hat{\rho}_b)=\Pi_k F_k(\hat{\rho}_a,\hat{\rho}_b)$ with [33]

$$F_k(\hat{\rho}_a,\hat{\rho}_b)=\frac{\mathrm{Tr}[e^{\vec{c}_k\hat{\vec{T}}_k}]}{\sqrt{\mathrm{Tr}[e^{\vec{a}_k\hat{\vec{T}}_k}]\mathrm{Tr}[e^{\vec{b}_k\hat{\vec{T}}_k}]}}=\frac{1+\sqrt{\frac{1}{2}(1+\cosh c_k)}}{\sqrt{\left(1+\cosh\frac{a_k}{2}\right)\left(1+\cosh\frac{b_k}{2}\right)}}, \tag{26}$$

$$\cosh c_k=\cosh(\beta^a E_k^a)\cosh(\beta^b E_k^b)\times\left\{1+\tanh(\beta^a E_k^a)\tanh(\beta^b E_k^b)\times\frac{\bar{\varepsilon}_k^a\bar{\varepsilon}_k^b+\mathrm{Re}[\Delta_k^a(\Delta_k^b)^*]}{E_k^a E_k^b}\right\}. \tag{27}$$

Here, we used the relation $\bar{\varepsilon}_k^a=\varepsilon_k^a-\mu^a$ and $a_k=2\beta^a E_k^a$, $E_k^a=\sqrt{(\bar{\varepsilon}_k^a)^2+|\Delta_k^a|^2}$, so that $\cosh(a_k/2)=\cosh(\beta^a E_k^a)$ [and analogously for $q_b=(V_b,T_b)$]. Note the explicit dependence of all the quantities on the temperature and the coupling strength, given through the superscripts a and b, denoting two parameter points $q_a=(T_a,V_a)$ and $q_b=(T_b,V_b)$.

Assuming that the chemical potential is also constant ($\mu=\varepsilon_F$) in the region of interest, where the phase transition takes place, the gap parameter reduces to $\Delta_k=\Delta$, for $|\bar{\varepsilon}|<\hbar\omega_D$ (and zero otherwise). Thus the self-consistent equation for the gap $\Delta_k=-\Sigma_{k'}V_{kk'}\{[1-2f(E_{k'})]/2E_{k'}\}\Delta_{k'}$ reads as $[1-2f(E_{k'})=\tanh(\beta E_{k'}/2)]$

$$1=D_F V\int_{-\hbar\omega_D}^{\hbar\omega_D}d\varepsilon\frac{\tanh\frac{\beta}{2}\sqrt{\varepsilon^2+\Delta^2(T,V)}}{2\sqrt{\varepsilon^2+\Delta^2(T,V)}}, \tag{28}$$

with $f(E)$ being the Fermi distribution.

In the $T\to 0$ limit, we obtain the expression for the ground state fidelity $F(|g_a\rangle\langle g_a|,|g_b\rangle\langle g_b|)=|\langle g_a|g_b\rangle|$. At zero temperature, the chemical potential is equal to the Fermi energy ε_F, $\mu(T=0)=\varepsilon_F$. Further, the gap equation reduces to $\Delta(V)=\hbar\omega_D/[\sinh(2/D_F V)]\simeq 2\hbar\omega_D\exp(-2/D_F V)$, where D_F is the density of states around the Fermi level. The zero temperature ground state fidelity is

$$F(|g_a\rangle\langle g_a|,|g_b\rangle\langle g_a|)=\prod_k\frac{1}{\sqrt{2}}\left(1+\frac{(\varepsilon_k-\varepsilon_F)^2+\Delta(V_a)\Delta(V_b)}{\sqrt{(\varepsilon_k-\varepsilon_F)^2+\Delta(V_a)^2}\sqrt{(\varepsilon_k-\varepsilon_F)^2+\Delta(V_b)^2}}\right)^{1/2},$$

which matches the expression one obtains for $T=0$. In other words, we see that, as in the case of the Stoner-Hubbard model, the point of criticality of the $T=0$ QPT can be inferred from the mixed state fidelity between the thermal states. The phase transition from superconductor to normal metal happens at $V=0$. Thus, for $V_a\to 0_+$ and $V_b=V_a+\delta V\to\delta V>0$, the fidelity between the corresponding ground states of Fermi sea $|g_F\rangle$ and the BCS superconductor $|g_{\mathrm{BCS}}\rangle$ becomes $|\langle g_F|g_{\mathrm{BCS}}\rangle|=\Pi_k(1/\sqrt{2})[1+|\varepsilon_k-\varepsilon_F|/\sqrt{(\varepsilon_k-\varepsilon_F)^2+\Delta(\delta V)^2}]^{1/2}$. We see that the BCS model at $T=0$ features the Anderson orthogonality catastrophe, just as the Stoner-Hubbard model.

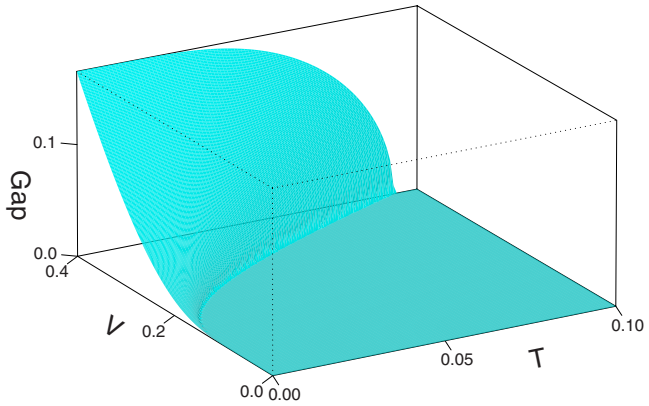


FIG. 5. (Color online) Gap $\Delta=\Delta(T,V)$ as a function of the temperature T and the coupling constant V. The plot is given in rescaled quantities $T\to k_BT/(\hbar\omega_D)$, $V\to D_FV$, and $\Delta\to\Delta/(\hbar\omega_D)$.

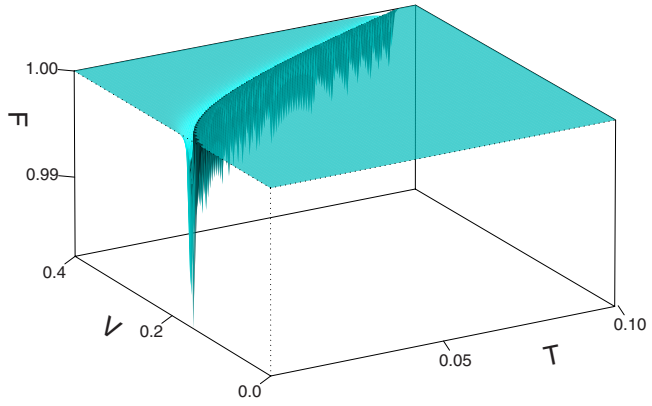


FIG. 7. (Color online) Fidelity $F=F(T,V)$ as a function of the temperature T and the coupling constant V. The plot is given in rescaled quantities $T\to k_BT/(\hbar\omega_D)$ and $V\to D_FV$, with $\delta T=0$ and $\delta V=10^{-3}$.

As in the previous case, in obtaining the numerical results for the fidelity, we used the subroutine hybrd.f from MINPACK [22]. Again, all the numerical results are given in rescaled quantities $T\to k_BT/(\hbar\omega_D)$, $V\to D_FV$, and $\Delta\to\Delta/(\hbar\omega_D)$. The result for the gap is given in Fig. 5. The line of the phase transition is clearly marked as the line along which the gap becomes nontrivial, and is presented in Fig. 6. Finally, the fidelity, with $\delta T=0$ and $\delta V=10^{-3}$, is plotted in Fig. 7. We varied the parameter differences δT and δV, and all the results obtained show the same qualitative picture—the fidelity exhibits a sudden drop from $F=1$ precisely along the line of the phase transition.

As already discussed, in the case of mutually commuting Hamiltonians, the fidelity reduces to the quantity

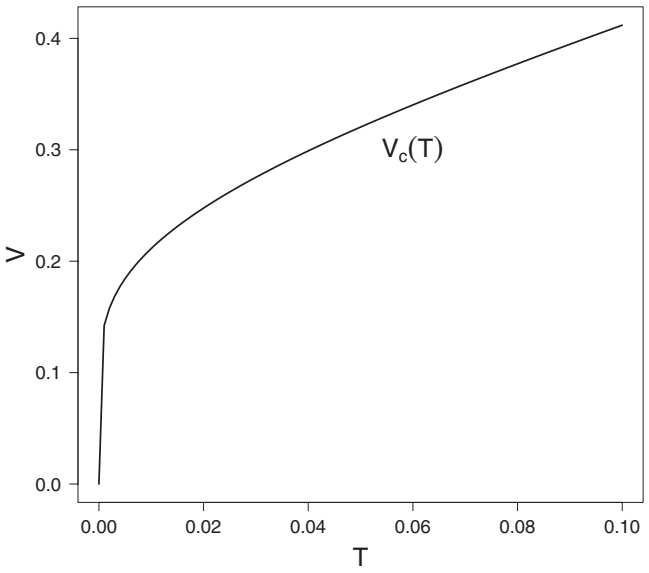


FIG. 6. Critical line $V_c=V_c(T)$ as a function of the temperature T. The plot is given in rescaled quantities $T\to k_BT/(\hbar\omega_D)$ and $V\to D_FV$.

$$C(\hat{\rho}_a,\hat{\rho}_b)=\frac{Z(\hat{\bar{H}})}{\sqrt{Z\left(\hat{\bar{H}}+\frac{\Delta\hat{H}}{2}\right)Z\left(\hat{\bar{H}}-\frac{\Delta\hat{H}}{2}\right)}} \tag{29}$$

which, through relation (18), establishes the connection between the singular behavior of the fidelity and the corresponding susceptibility (or the heat capacity, etc.). In the noncommuting case, the same relation between $C(\hat{\rho}_a,\hat{\rho}_b)$ and χ is still valid (see Appendix D), yet the fidelity is in general *not* identically equal to C.

If we approach a line of the phase transition along a curve $q=q(\alpha)$, $\alpha\in\mathbb{R}$, we find $F=C(1+\mathcal{F}d\alpha^2)$, where $d\alpha$ defines the difference δq. For reasons of simplicity, we omit here the explicit, lengthy expression for $\mathcal{F}$. Although relatively complicated and difficult to study directly, it is evident that, apart from the Hamiltonian's eigenvalues, it is also explicitly given by the rate of change, with respect to α, of the Hamiltonian's eigenbasis. This is also evident from the fact that for the case of mutually commuting Hamiltonians, when the eigenbasis is common, $C\equiv F$ and thus $\mathcal{F}\equiv 0$. In the commuting case, the change of state is given by the change of the Hamiltonian's eigenvalues only, while in the noncommuting case, the overall change of state is given by the change of *both* the eigenvalues and the eigenvectors. Thus, the singularity of $\mathcal{F}$ can on its own mark the drastic change in the structure of the system's eigenbasis, thus bringing about the finite difference between C and F. This is indeed the case for BCS superconductors, where the difference $C-F$ becomes nontrivial precisely along the line of the phase transition, where $C<F$; see Fig. 8. We see that, as intuitively expected, the state of a system exhibits an abrupt change in its structure along the line of the phase transition both in terms of its eigenvalues, as well as in terms of its eigenstates [for the structural analysis of the system's eigenstates given by a parametrized Hamiltonian $\hat{H}(q)$, see for example [23,24]].

Another way to quantify the structural change of the eigenvectors is through the Uhlmann connection and the corresponding mixed-state geometric phase [19], the mixed-

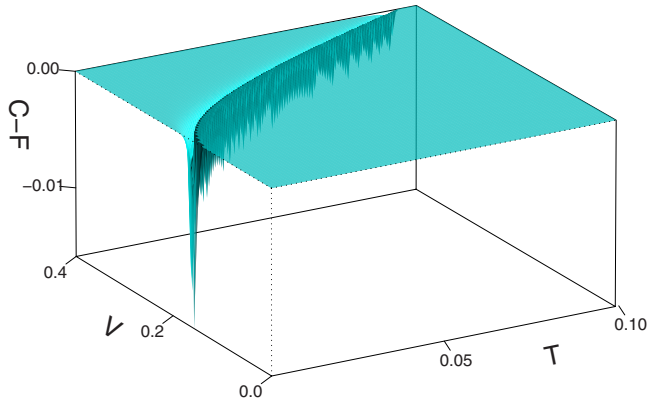


FIG. 8. (Color online) Difference $C(T,V)-F(T,V)$ as a function of the temperature T and the coupling constant V. The plot is given in rescaled quantities $T\rightarrow k_BT/(\hbar\omega_D)$ and $V\rightarrow D_FV$, with $\delta T=0$ and $\delta V=10^{-3}$.

state generalization of the Berry connection and phase. The fidelity $F(\hat{\rho}_a,\hat{\rho}_b)=\mathrm{Tr}\sqrt{\sqrt{\hat{\rho}_a}\hat{\rho}_b\sqrt{\hat{\rho}_a}}$ can also be expressed as $F(\hat{\rho}_a,\hat{\rho}_b)=\mathrm{Tr}[|\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}|]$, where $|\hat{A}|=(\hat{A}\hat{A}^\dagger)^{1/2}$ represents the modulus of the operator $\hat{A}$ (see [25]). The Uhlmann parallel transport condition (i.e., the connection) is given by the unitary operator $\hat{U}_{ab}$, such that $F(\hat{\rho}_a,\hat{\rho}_b)=\mathrm{Tr}[\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}\hat{U}_{ab}]$ [see Eq. (9) from [25]]. In other words, the connection operator defining the parallel transport is the inverse of the unitary $\hat{V}_{ab}$, given by the polar decomposition [1] $\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}=|\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}|\hat{V}_{ab}$, i.e., $\hat{U}_{ab}=\hat{V}_{ab}^\dagger$. Note that the parallel transport condition (the connection), given by $\hat{U}_{ab}$, induces both the local Uhlmann curvature two-form, as well as the global mixed-state geometric phase [see Eqs. (11) and (12) from [25]]. Let us define

$$H(\hat{\rho}_a,\hat{\rho}_b)=\mathrm{Tr}[\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}]. \tag{30}$$

Obviously, in the case of mutually commuting Hamiltonians $H=F(=C)$ and the Uhlmann connection is trivial, $\hat{U}_{ab}=\hat{I}$. In the case of the BCS model, we have evaluated the quantity $H-F$, and the result is qualitatively the same as for $C-F$. In other words, along the line of the phase transition we have found the strict inequality $H-F=\mathrm{Tr}[|\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}|(\hat{U}_{ab}-\hat{I})]<\mathrm{const}<0$. From the Cauchy-Schwarz inequality it follows then that $\hat{U}_{ab}-\hat{I}\neq 0$, or $\hat{U}_{ab}\neq\hat{I}$—a clearly abrupt change of the connection operator $\hat{U}_{ab}$ occurs and it becomes nontrivial in the vicinity of the line of the phase transition. Since the two parameter points q_a and q_b are taken to be close to each other, such a behavior implies the nonanalyticity of the local Uhlmann curvature form along the line of the phase transition, which in turn results in generally nontrivial global Uhlmann mixed-state geometric phase (see, for example, the discussions in [13,14]). This can be seen as a mixed thermal state generalization of recent results [15] on the relation between QPTs and Berry geometric phases. Further, we have that along the line of the phase transition we have $C<H<F$, while $C\simeq H\simeq F\simeq 1$ otherwise.

IV. CONCLUSION

In this paper we analyzed the fidelity approach to both zero temperature (quantum) as well as finite temperature phase transitions. It is based on the notion of quantum state distinguishability, applied to the case of macroscopic many-body systems whose states determine the global order of the system and its phase. We focused on the two particular cases of the Stoner-Hubbard model for itinerant electron magnetism and the BCS theory of superconductivity, as representatives of two distinct classes of physical systems: Those defined by a set of mutually commuting Hamiltonians and those defined by mutually *non*commuting Hamiltonians, with respect to a given parameter space. We found that in both cases the fidelity can mark the regions of PTs by its sudden drop in the vicinity of the transition line. We discussed in detail the general case of mutually commuting Hamiltonians, where in the case of symmetry-breaking LGW type of transitions the nonanalyticity of fidelity is a direct consequence of the nonanalyticity of the corresponding "generalized" susceptibility. The case of mutually noncommuting Hamiltonians is more complex as there the structure of the Hamiltonian's eigenvectors directly affects the system's thermal state thus introducing a feature relevant for the macroscopic phase distinguishability. On the example of the BCS superconductivity we showed that the phase transition is accompanied by the abrupt change of the Hamiltonian's eigenvectors, quantified by the differences $C-F$ and $H-F$. The second one is of particular interest as it determines the Uhlmann connection and the associated mixed-state geometric phase, which generically becomes nontrivial in the vicinity of the phase transition only—a mixed state generalization of the emergence of nontrivial Berry geometric phase in the vicinity of criticality for QPTs.

There are two main extensions of this work. First, the study of more general phase transitions that fall outside the standard symmetry-breaking LGW paradigm. Second, the general study of the structural analysis of the system's eigenstates in the case of mutually noncommuting Hamiltonians, given within the framework of the mixed state fidelity and the Uhlmann geometric phase.

ACKNOWLEDGMENTS

The authors thank EU FEDER and FCT projects Quant-Log POCI/MAT/55796/2004 and QSec PTDC/EIA/67661/2006. N.P. acknowledges the support from EU FEDER and FCT grant SFRH/BPD/31807/2006. V.R.V. acknowledges the support of projects SFA-2-91, POCI/FIS/58746/2004, and FCT PTDC/FIS/70843/2006 and also useful discussions with Miguel A. N. Araújo and Pedro D. Sacramento.

APPENDIX A

In this appendix we prove various technical statements relevant for the analysis of the ground state fidelity in the Stoner-Hubbard model.

1. Proving the relation $\varepsilon_F=\tilde{\varepsilon}_{k_{F\uparrow}}=\tilde{\varepsilon}_{k_{F\downarrow}}$

Using $N_\sigma=\int_0^{N_\sigma}dN_\sigma$, the ground state energy can be written as

$$E_g = \int_0^{N_\uparrow} E_{k\uparrow} dN_\uparrow + \int_0^{N_\downarrow} E_{k\downarrow} dN_\downarrow - U\frac{N_\uparrow N_\downarrow}{V}$$
$$= \int_0^{N_\uparrow} \varepsilon_k dN_\uparrow + \int_0^{N_\downarrow} \varepsilon_k dN_\downarrow + U\frac{N_\uparrow N_\downarrow}{V}. \quad \text{(A1)}$$

Since it is the local minima, for given $N=N_\uparrow+N_\downarrow$ we have that

$$\frac{\partial(E_g-\mu N)}{\partial N_\uparrow} = \varepsilon_{k_F\uparrow} - \mu + U\frac{N_\downarrow}{V} = 0,$$

$$\frac{\partial(E_g-\mu N)}{\partial N_\downarrow} = \varepsilon_{k_F\downarrow} - \mu + U\frac{N_\uparrow}{V} = 0, \quad \text{(A2)}$$

and we immediately get $E_{k_{F\uparrow}}=\varepsilon_{k_F\uparrow}+Un_\downarrow=\varepsilon_{k_F\downarrow}+Un_\uparrow=E_{k_{F\downarrow}}=\varepsilon_F$, with $\mu=\varepsilon_F$, for $T=0$.

2. Deriving U_c from Stoner criterion

The density of states is defined as $D(\varepsilon)=(1/V)(dN_\sigma/d\varepsilon)$ in the paramagnetic phase, where ε is the one-particle energy and $N_\sigma(\varepsilon)$ is the number of states for each spin direction whose energy is smaller or equal than ε. For each spin projection, we have $N_\sigma/V=k_F^3/(6\pi^2)$, while from $\varepsilon(k)=(\hbar^2/2m)k^2$ we have that $k=[(2m/\hbar^2)\varepsilon]^{1/2}$. Thus $(N_\sigma(\varepsilon)/V)=(1/6\pi^2)[(2m/\hbar^2)\varepsilon]^{3/2}$ and

$$D(\varepsilon) = \frac{1}{V}\frac{dN_\sigma(\varepsilon)}{d\varepsilon} = \frac{3}{2}\frac{1}{6\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{\varepsilon} = \frac{3}{2\varepsilon}\frac{N_\sigma}{V} = \frac{3}{4\varepsilon}\frac{N}{V}. \quad \text{(A3)}$$

The difference between the up and down single particle energies $E_{k\uparrow}$ and $E_{k\downarrow}$ defines the self-consistent magnetic field $\Delta=2U[M(\Delta)/V]$. The phase transition occurs when Δ becomes nonzero, leading to the equation $1=(2U_c/V)\times(\partial M/\partial\Delta)|_{\Delta=0}$, which also gives the vanishing of the susceptibility denominator in the random phase approximation (Stoner enhancement factor) [20]. One has then $U_cD_F=1$, and therefore we find $U_c=1/D_F=\frac{4}{3}\varepsilon_F(V/N)$.

3. Deriving U_c from Eqs. (9) and (10)

Using $x=k_{F\uparrow}$, $y=k_{F\downarrow}$, Eq. (9) can be rewritten as $f(x,y)\equiv x^3+y^3-a=0$, with $a=6\pi^2(N/V)\equiv 2k_F^3$. Also, the quadratic term from Eq. (10) that is for $U>U_c$ responsible for nontrivial magnetic solutions reads as $g(x,y;U)\equiv(x+y)-(U/\alpha)(x^2+xy+y^2)=0$, with $\alpha=3\pi^2(\hbar^2/m)\equiv\frac{3}{2}k_FU_c$.

First, we show that for $U>U_c$ [see Eq. (8)], curves f and g have intersection in two points, symmetric with respect to the $y=x$ line, while for $U=U_c$ they touch precisely in the point $(x,y)=(k_F,k_F)$. In the coordinate system $(\bar{x},\bar{y})$, rotated by the angle $\varphi=\pi/4$ and translated by the vector $(A,0)$ from the system (x,y), the curve $g(\bar{x},\bar{y};U)=0$ becomes [34]

$$\frac{\bar{x}^2}{A^2}+\frac{\bar{y}^2}{B^2}=1, \quad \text{(A4)}$$

with $A=(\sqrt{2}/3)(k_F/U)$ and $B=\sqrt{2/3}(k_F/U)$—a real ellipse with the main axes $A=A(U)$ and $B=B(U)$. As the parameter U increases from 0 (when $A,B\to\infty$), the main axes A and B decrease and eventually the ellipse g touches the curve f in the point $(\bar{x},\bar{y})=(2A,0)$, which in the original coordinate system correspond to the point $(x,y)=(k_F,k_F)$. In other words, $2A=\sqrt{2}k_F$ and $U=U_c$. Further increase of U beyond U_c results in even smaller ellipses that intersect the curve f in two symmetric points (note that the curve f is itself symmetric along the $y=x$ line).

To show that for $U>U_c$ the "balanced" solutions of Eqs. (9) and (10) are indeed the physical ones, we calculate the total ground state energy E_g of the system:

$$E_g = V\int_{E_{0\uparrow}}^{E_{F\uparrow}} \tilde{\varepsilon}_{k\uparrow} D(E_{k\uparrow}) dE_{k\uparrow} + V\int_{E_{0\downarrow}}^{E_{F\downarrow}} E_{k\downarrow} D(E_{k\downarrow}) dE_{k\downarrow} - VUn_\uparrow n_\downarrow. \quad \text{(A5)}$$

Upon evaluating the above integrals, using $M=(N_\uparrow-N_\downarrow)/2$, the total ground state energy reads

$$E_g = \frac{V}{20\pi^2}\frac{\hbar^2}{m}\left(3\pi^2\frac{N}{V}\right)^{5/3}\left[\left(1+\frac{2M}{N}\right)^{5/3}+\left(1-\frac{2M}{N}\right)^{5/3}\right] + \frac{UN^2}{4V}\left[1-\left(\frac{2M}{N}\right)^2\right]. \quad \text{(A6)}$$

From the above expression, we see that for every value of the coupling constant U, $(\partial E_g/\partial M)(U,M=0)=0$, while $(\partial^2E_g/\partial M^2)(U,M=0)$ changes the sign from positive to negative precisely in $U=U_c$. Thus we indeed have that for $U>U_c$ the "balanced" magnetic solutions with $M\neq 0$ have lower energy.

4. Deriving the maximum fidelity, i.e., the maximum derivative dk_F/dU for $U=U_c$, from Eqs. (9) and (10)

From $df=0$ and $dg=0$, we get ($\partial f/\partial x=f_x$, $dx/dU=\dot{x}$, etc.)

$$f_x\dot{x}+f_y\dot{y}=0,$$
$$g_x\dot{x}+g_y\dot{y}=-g_u. \quad \text{(A7)}$$

Solving it for $\dot{x}$ and $\dot{y}$, we obtain

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \frac{g_u}{f_xg_y-f_yg_x}\begin{bmatrix}f_y\\ -f_x\end{bmatrix}. \quad \text{(A8)}$$

Finally,

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \frac{3}{4}\frac{k_f}{D_FV^2}\frac{x+y}{x-y}\frac{1}{xy}\begin{bmatrix}y^2\\ -x^2\end{bmatrix}. \quad \text{(A9)}$$

Thus, in the limit $U\to U_{c+}$ [note that $g(x,y;U)=0$ is valid for $U\geq U_c$], we have that $x\to k_{F+}$, $y\to k_{F-}$ and $\dot{x}\to+\infty$, $\dot{y}\to-\infty$ (or vice versa, for the $x<y$ solution branch). In other words, $dV(k_{F\downarrow})/dU=4\pi(k_{F\downarrow})^2[d(k_{F\downarrow})/dU]=4\pi x^2\dot{x}\to+\infty$ and $dV(k_{F\uparrow})/dU\to-\infty$, or vice versa.

APPENDIX B

Below, the basic identities of the su(2) algebra of the electron spin operators leading to Eq. (15) are proven. We dis-

cuss operators having a fixed momentum and drop the index k for convenience.

1. $\hat{\vec{S}}$-Algebra identities

In general, operators defined by $\hat{S}^a=\hat{c}^\dagger_\alpha S^a_{\alpha\beta}\hat{c}_\beta$, where $\hat{c}^\dagger_\alpha$ and $\hat{c}_\beta$ are either bosonic or fermionic operators and $S^a_{\alpha\beta}$ is a matrix representation of a spin S algebra, i.e., satisfying $[S^a,S^b]_{\alpha\beta}=i\varepsilon^{abc}S^c_{\alpha\beta}$, define a su(2) algebra. This is the case of the electron spin operators defined by $\hat{\vec{S}}=\hat{\Psi}^\dagger(\vec{\sigma}/2)\hat{\Psi}$, where $\hat{\Psi}^\dagger=[\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow]$ and $\vec{\sigma}$ is the vector of Pauli matrices, and given by $\hat{S}^z=\frac{1}{2}(\hat{n}_\uparrow-\hat{n}_\downarrow)$, $\hat{S}^+=\hat{c}^\dagger_\uparrow\hat{c}_\downarrow$ and $\hat{S}^-=\hat{c}^\dagger_\downarrow\hat{c}_\uparrow$ (with $\hat{S}^\pm=\hat{S}^x\pm i\hat{S}^y$).

Using the anticommutation relations for the one-electron modes $\hat{c}^\dagger_\sigma$, one can easily obtain the following commutation and anticommutation identities:

$$[\hat{S}^z,\hat{S}^\pm]=\pm\hat{S}^\pm,\quad [\hat{S}^+,\hat{S}^-]=2\hat{S}^z,$$

$$\{\hat{S}^z,\hat{S}^\pm\}=0,\quad \{\hat{S}^+,\hat{S}^-\}=\hat{I}_s, \tag{B1}$$

with $\hat{I}_s\equiv\hat{n}_\uparrow+\hat{n}_\downarrow-2\hat{n}_\uparrow\hat{n}_\downarrow=(2-\hat{n})\hat{n}$, $\hat{n}=\hat{n}_\uparrow+\hat{n}_\downarrow$, the projector onto the subspace of the single-occupied states being an identity operator for this algebra, since $\hat{I}_s^2=\hat{I}_s$, $\hat{I}_s\hat{\vec{S}}=\hat{\vec{S}}\hat{I}_s=\hat{\vec{S}}$ and $(\hat{\vec{S}})^2=\frac{3}{4}\hat{I}_s$. One of the immediate consequences of the above commutation relations is that $\{\hat{S}^x,\hat{S}^y,\hat{S}^z\}$ form a su(2) algebra.

Further, we can obtain the traces $\mathrm{Tr}[\hat{I}]=4$, $\mathrm{Tr}[\hat{I}_s]=2$ and $\mathrm{Tr}[\hat{\vec{S}}]=\vec{0}$.

2. Evaluating $e^{\vec{h}\hat{\vec{S}}}$ and $\mathrm{Tr}[e^{\vec{h}\hat{\vec{S}}}]$

For $e^{\vec{h}\hat{\vec{S}}}$, we have

$$e^{\vec{h}\hat{\vec{S}}}=\sum_{n=0}^{+\infty}\frac{1}{n!}(\vec{h}\hat{\vec{S}})^n=\hat{I}+\sum_{n=1}^{+\infty}\frac{1}{n!}(\vec{h}\hat{\vec{S}})^n. \tag{B2}$$

For $(\vec{h}\hat{\vec{S}})^2$, we have

$$\begin{aligned}(\vec{h}\hat{\vec{S}})^2&=\left[\frac{1}{2}(h^+\hat{S}^-+h^-\hat{S}^+)+h_z\hat{S}^z\right]^2\\&=\frac{1}{4}(h^+)^2(\hat{S}^-)^2+\frac{1}{4}(h^-)^2(\hat{S}^+)^2+h_z^2(\hat{S}^z)^2\\&\quad+\frac{1}{4}(h^+h^-)\{\hat{S}^-,\hat{S}^+\}+\frac{1}{2}(h_zh^+)\{\hat{S}^z,\hat{S}^-\}\\&\quad+\frac{1}{2}(h_zh^-)\{\hat{S}^z,\hat{S}^+\}.\end{aligned} \tag{B3}$$

Using the $\hat{\vec{S}}$-algebra identities (B1), we get

$$(\vec{h}\hat{\vec{S}})^2=\frac{1}{4}(h_z)^2\hat{I}_s+\frac{1}{4}(h^+h^-)\hat{I}_s=\frac{1}{4}|h|^2\hat{I}_s, \tag{B4}$$

where [35] $|h|^2=(h_z)^2+h^+h^-\equiv h^2$. Thus we obtain

$$\begin{aligned}e^{\vec{h}\hat{\vec{S}}}&=\hat{I}+\sum_{n=1}^{+\infty}\frac{1}{(2n)!}\left(\frac{h}{2}\right)^{2n}\hat{I}_s+\sum_{n=0}^{+\infty}\frac{1}{(2n+1)!}\left(\frac{h}{2}\right)^{2n}(\vec{h}\hat{\vec{S}})\\&=(\hat{I}-\hat{I}_s)+\left[\sum_{n=0}^{+\infty}\frac{1}{(2n)!}\left(\frac{h}{2}\right)^{2n}\right]\hat{I}_s\\&\quad+\left[\sum_{n=0}^{+\infty}\frac{1}{(2n+1)!}\left(\frac{h}{2}\right)^{2n+1}\right]\frac{2}{h}(\vec{h}\hat{\vec{S}}),\end{aligned}$$

$$e^{\vec{h}\hat{\vec{S}}}=(\hat{I}-\hat{I}_s)+\cosh\left(\frac{h}{2}\right)\hat{I}_s+2\sinh\left(\frac{h}{2}\right)\frac{(\vec{h}\hat{\vec{S}})}{h}. \tag{B5}$$

Using the trace formulas, we finally obtain

$$\mathrm{Tr}[e^{\vec{h}\hat{\vec{S}}}]=2\left(1+\cosh\frac{h}{2}\right). \tag{B6}$$

3. Evaluating $e^{\alpha\hat{n}}$ and $\mathrm{Tr}[e^{\alpha\hat{n}}]$

Enlarging the su(2) algebra of the $\hat{\vec{S}}$ operators to the operator $\hat{n}=\hat{n}_\uparrow+\hat{n}_\downarrow$ it is useful to know its algebraic properties and their relation to the other operators. We first evaluate $\hat{n}^k$. Since $\hat{n}_\sigma^2=\hat{n}_\sigma$ and $\hat{n}^2=(\hat{n}_\uparrow+\hat{n}_\downarrow)^2=\hat{n}+2\hat{n}_\uparrow\hat{n}_\downarrow$ we find that $\hat{n}^3=-2\hat{n}+3\hat{n}^2$ and in general $\hat{n}^k=-a_k\hat{n}+b_k\hat{n}^2$, where a_k and b_k satisfy the recurrence equations $a_{k+1}=2b_k$ and $b_{k+1}=3b_k-a_k$, leading to $\hat{n}^k=-(2^{k-1}-2)\hat{n}+(2^{k-1}-1)\hat{n}^2$. The exponential $e^{\alpha\hat{n}}$ is given by

$$\begin{aligned}e^{\alpha\hat{n}}&=\hat{I}-\left(\frac{e^{2\alpha}}{2}-2e^\alpha+\frac{3}{2}\right)\hat{n}+\left(\frac{e^{2\alpha}}{2}-e^\alpha+\frac{1}{2}\right)\hat{n}^2\\&=\hat{I}+u\hat{n}+v\hat{n}^2,\end{aligned} \tag{B7}$$

which can be easily verified since $\hat{n}=\hat{n}_\uparrow+\hat{n}_\downarrow$ can only take the values $n=0$, 1, and 2.

Using the obvious trace relations $\mathrm{Tr}[\hat{n}]=4$ and $\mathrm{Tr}[\hat{n}^2]=6$, we also get $\mathrm{Tr}[e^{\alpha\hat{n}}]=4+4u+6v$, or

$$\mathrm{Tr}[e^{\alpha\hat{n}}]=(e^\alpha+1)^2. \tag{B8}$$

The number operator commutes with the $\hat{\vec{S}}$ operators, and one has $\hat{n}\hat{I}_s=\hat{I}_s\hat{n}=\hat{I}_s$ and $\hat{n}\hat{\vec{S}}=\hat{\vec{S}}\hat{n}=\hat{\vec{S}}$. Finally, one obtains

$$\begin{aligned}\mathrm{Tr}[e^{(\alpha\hat{n}+\vec{h}\hat{\vec{S}})}]&=(1+e^{\alpha+h/2})(1+e^{\alpha-h/2})\\&=2e^\alpha\left(\cosh\alpha+\cosh\frac{h}{2}\right).\end{aligned} \tag{B9}$$

APPENDIX C

In the following, we obtain the basic formulas for the su(2) algebra of the Nambu operators $\hat{\vec{T}}$ used in evaluating Eq. (26) for the fidelity between two states of a BCS superconductor in thermodynamical equilibrium.

1. $\hat{\vec{T}}$-Algebra identities

We now consider the operators defined by $\hat{\vec{T}}=\hat{\psi}^\dagger(\vec{\sigma}/2)\hat{\psi}$, where $\hat{\psi}^\dagger=[\hat{c}_\uparrow^\dagger \hat{c}_\downarrow]$ and $\vec{\sigma}$ is the vector of Pauli matrices, and given by $\hat{T}^z=\frac{1}{2}(\hat{n}_\uparrow+\hat{n}_\downarrow-1)$, $\hat{T}^+=\hat{c}_\uparrow^\dagger\hat{c}_\downarrow^\dagger$, and $\hat{T}^-=\hat{c}_\downarrow\hat{c}_\uparrow$ (with $\hat{T}^\pm=\hat{T}^x\pm i\hat{T}^y$). These operators $\hat{T}^a$ are of the form considered before, since the fermion anticommutation relations are invariant under the electron-hole transformation, i.e., under the interchange $\hat{c}\leftrightarrow\hat{c}^\dagger$. Similarly to the previous case of the $\hat{\vec{S}}$ algebra, we obtain the commutation and anticommutation relations:

$$[\hat{T}^0,\hat{T}^\pm]=\pm\hat{T}^\pm,\quad [\hat{T}^+,\hat{T}^-]=2\hat{T}^0,$$

$$\{\hat{T}^0,\hat{T}^\pm\}=0,\quad \{\hat{T}^+,\hat{T}^-\}=\hat{I}_t, \tag{C1}$$

with $\hat{I}_t\equiv 2\hat{n}_\uparrow\hat{n}_\downarrow-(\hat{n}_\uparrow+\hat{n}_\downarrow)+1=(n-1)^2$, the projector onto the subspace of the empty and doubly occupied states being an identity operator for this algebra, since $\hat{I}_t^2=\hat{I}_t$, $\hat{I}_t\hat{\vec{T}}=\hat{\vec{T}}\hat{I}_t=\hat{\vec{T}}$ and $(\hat{\vec{T}})^2=\frac{3}{4}\hat{I}_t$. One of the immediate consequences of the above commutation relations is that $\{\hat{T}^0,\hat{T}^+,\hat{T}^-\}$ form a su(2) algebra. Further, simple algebra gives

$$\mathrm{Tr}[\hat{\vec{T}}]=0,\quad \mathrm{Tr}[\hat{I}_t]=2. \tag{C2}$$

Note that in the above definition of $\hat{\psi}^\dagger$, the operators $\hat{c}_\uparrow^\dagger$ and $\hat{c}_\downarrow$ have the opposite momenta, k and $-k$, respectively. Had the momenta been defined to be the same, the two sets of operators $\hat{\vec{S}}$ and $\hat{\vec{T}}$ would commute with each other, i.e., $[\hat{S}^a,\hat{T}^b]=0$. Also, one would have $\hat{I}_s+\hat{I}_t=\hat{I}$.

2. Evaluating $e^{\vec{a}\hat{\vec{T}}}$ and $\mathrm{Tr}[e^{\vec{a}\hat{\vec{T}}}]$

For $e^{\vec{a}\hat{\vec{T}}}$, we have

$$e^{\vec{a}\hat{\vec{T}}}=\sum_{n=0}^{+\infty}\frac{1}{n!}(\vec{a}\hat{\vec{T}})^n=\hat{I}+\sum_{n=1}^{+\infty}\frac{1}{n!}(\vec{a}\hat{\vec{T}})^n. \tag{C3}$$

Similarly to the previous case, for $(\vec{a}\hat{\vec{T}})^2$, we have

$$(\vec{a}\hat{\vec{T}})^2=\left[a_z\hat{T}^0+\frac{1}{2}(a^+\hat{T}^-+a^-\hat{T}^+)\right]^2=\frac{1}{4}|a|^2\hat{I}_t, \tag{C4}$$

where $|a|^2=(a_z)^2+a^+a^-\equiv a^2$. Thus we obtain ($\hat{Z}_t=\hat{I}-\hat{I}_t$):

$$\begin{aligned}e^{\vec{a}\hat{\vec{T}}}&=\hat{I}+\sum_{n=1}^{+\infty}\frac{1}{(2n)!}\left(\frac{a}{2}\right)^{2n}\hat{I}_t+\sum_{n=0}^{+\infty}\frac{1}{(2n+1)!}\left(\frac{a}{2}\right)^{2n}(\vec{a}\hat{\vec{T}})\\&=(\hat{I}-\hat{I}_t)+\left[\sum_{n=0}^{+\infty}\frac{1}{(2n)!}\left(\frac{a}{2}\right)^{2n}\right]\hat{I}_t\\&\quad+\left[\sum_{n=0}^{+\infty}\frac{1}{(2n+1)!}\left(\frac{a}{2}\right)^{2n+1}\right]\frac{2}{a}(\vec{a}\hat{\vec{T}}),\end{aligned}$$

$$e^{\vec{a}\hat{\vec{T}}}=\hat{Z}_t+\cosh\left(\frac{a}{2}\right)\hat{I}_t+2\sinh\left(\frac{a}{2}\right)\frac{(\vec{a}\hat{\vec{T}})}{a}. \tag{C5}$$

Using the trace formulas (C2), we finally obtain

$$\mathrm{Tr}[e^{\vec{a}\hat{\vec{T}}}]=2\left(1+\cosh\frac{a}{2}\right). \tag{C6}$$

3. Evaluating $e^{2\vec{c}\hat{\vec{T}}}=e^{(\vec{a}/2)\hat{\vec{T}}}e^{\vec{b}\hat{\vec{T}}}e^{(\vec{a}/2)\hat{\vec{T}}}$ and $\mathrm{Tr}[e^{\vec{c}\hat{\vec{T}}}]$

From the above formula (C6) and the known trigonometric relation $[\cosh(\alpha/2)]^2=\frac{1}{2}(1+\cosh\alpha)$, we have that $\mathrm{Tr}[e^{\vec{c}\hat{\vec{T}}}]=2[1+\cosh(c/2)]=2\left[1+\sqrt{\frac{1}{2}(1+\cosh c)}\right]$. Thus, by evaluating $e^{2\vec{c}\hat{\vec{T}}}$, we obtain the result for $\cosh c$ and therefore $\mathrm{Tr}[e^{\vec{c}\hat{\vec{T}}}]$. First, we obtain the general expression for $e^{\vec{a}\hat{\vec{T}}}e^{\vec{b}\hat{\vec{T}}}$. Using the result (C5) for $e^{\vec{a}\hat{\vec{T}}}$ (and analogously for $e^{\vec{b}\hat{\vec{T}}}$), we get

$$\begin{aligned}e^{\vec{a}\hat{\vec{T}}}e^{\vec{b}\hat{\vec{T}}}&=\hat{Z}_t+\left(\cosh\frac{a}{2}\cosh\frac{b}{2}\right)\hat{I}_t+2\left[\left(\cosh\frac{a}{2}\sinh\frac{b}{2}\right)\frac{\vec{b}}{b}\right.\\&\quad\left.+\left(\sinh\frac{a}{2}\cosh\frac{b}{2}\frac{\vec{a}}{a}\right)\right]\hat{\vec{T}}+4\left(\sinh\frac{a}{2}\sinh\frac{b}{2}\right)\frac{\vec{a}\hat{\vec{T}}}{a}\frac{\vec{b}\hat{\vec{T}}}{b}.\end{aligned} \tag{C7}$$

In deriving the above expression, we have used the above identities (C2). As in Eq. (C4), using Eq. (C1) we obtain

$$(\vec{a}\hat{\vec{T}})(\vec{b}\hat{\vec{T}})=\frac{1}{4}[(\vec{a}\vec{b})\hat{I}_t+2i(\vec{a}\times\vec{b})\hat{\vec{T}}]. \tag{C8}$$

Therefore, we have

$$\begin{aligned}e^{\vec{a}\hat{\vec{T}}}e^{\vec{b}\hat{\vec{T}}}&=\hat{Z}_t+\left[\left(\cosh\frac{a}{2}\cosh\frac{b}{2}\right)+\left(\sinh\frac{a}{2}\sinh\frac{b}{2}\right)\frac{(\vec{a}\vec{b})}{ab}\right]\hat{I}_t\\&\quad+2\left[\left(\sinh\frac{a}{2}\cosh\frac{b}{2}\right)\frac{\vec{a}}{a}+\left(\cosh\frac{a}{2}\sinh\frac{b}{2}\right)\frac{\vec{b}}{b}\right.\\&\quad\left.+i\frac{(\vec{a}\times\vec{b})}{ab}\right]\hat{\vec{T}}.\end{aligned} \tag{C9}$$

Applying the above result twice, we finally obtain the expression for $e^{2\vec{c}\hat{\vec{T}}}$:

$$\begin{aligned}e^{2\vec{c}\hat{\vec{T}}}&=e^{(\vec{a}/2)\hat{\vec{T}}}e^{\vec{b}\hat{\vec{T}}}e^{(\vec{a}/2)\hat{\vec{T}}}=\hat{Z}_t+\left[\left(\cosh\frac{a}{2}\cosh\frac{b}{2}\right)\right.\\&\quad\left.+\left(\sinh\frac{a}{2}\sinh\frac{b}{2}\right)\frac{(\vec{a}\vec{b})}{ab}\right]\hat{I}_t+\left[\left(\sinh\frac{a}{2}\cosh\frac{b}{2}\right)\frac{\vec{a}}{a}\right.\\&\quad\left.+2\left(\sinh\frac{b}{2}\right)^2\frac{\vec{b}}{b}+\left(\cosh\frac{a}{2}-1\right)\left(\sinh\frac{b}{2}\right)\frac{(\vec{a}\vec{b})\vec{a}}{a^2b}\right]\hat{\vec{T}}.\end{aligned} \tag{C10}$$

Comparing the above result with the expression (C5), we eventually end up with the expression for $\cosh c$:

$$\cosh c = \left(\cosh\frac{a}{2}\cosh\frac{b}{2}\right) + \left(\sinh\frac{a}{2}\sinh\frac{b}{2}\right)\frac{(\vec{a}\vec{b})}{ab} \tag{C11}$$

which reduces to Eq. (27) for particular values of $\vec{a}_k = \tilde{\vec{h}}_k(q_a)$ and $\vec{b}_k = \tilde{\vec{h}}_k(q_b)$.

APPENDIX D

In this appendix we prove that in the case of mutually noncommuting Hamiltonians a relation analogous to Eq. (18) holds between C, given by Eq. (29), and the susceptibility χ. As in the commuting case, for simplicity, we consider a Hamiltonian $\hat{H}=\hat{H}_0-h\hat{S}$ with the symmetry-breaking term $\hat{S}$, and $h=h(q)$. Note that, in this case, the two terms in the Hamiltonian *do not commute* with each other, $[\hat{H}_0,\hat{S}]\neq 0$. Thus we have the following imaginary time Dyson expansion around the point $h=0$ [36]:

$$e^{-\beta(\hat{H}_0-h\hat{S})} \simeq \left\{ e^{-\beta\hat{H}_0} + h\int_0^\beta d\tau\, e^{-\beta\hat{H}_0}\hat{S}(\tau) + h^2\int_0^\beta d\tau\int_0^\tau d\tau_1 e^{-\beta\hat{H}_0}\hat{S}(\tau)\hat{S}(\tau_1)\right\}, \tag{D1}$$

with $\hat{S}(\tau)=e^{\tau\hat{H}_0}\hat{S}e^{-\tau\hat{H}_0}$. From the above equation, we obtain the expressions for the magnetization $M=\langle\hat{S}\rangle$ and the susceptibility $\chi=\partial M/\partial h$ given by derivatives of the partition function Z. First, the magnetization can be expressed as (using the commutativity between the partial derivative and the trace, $(\partial/\partial h)\mathrm{Tr}[\cdot]=\mathrm{Tr}(\partial/\partial h)[\cdot]$)

$$\begin{aligned} M &= \frac{1}{\beta}\frac{\partial \ln Z}{\partial h} = \frac{1}{\beta}\frac{1}{Z}\frac{\partial Z}{\partial h} = \frac{1}{\beta}\frac{1}{Z}\frac{\partial}{\partial h}\mathrm{Tr}[e^{-\beta\hat{H}}] \\ &= \frac{1}{\beta}\frac{1}{Z}\int_0^\beta d\tau\,\mathrm{Tr}[e^{-\beta\hat{H}_0}e^{\tau\hat{H}_0}\hat{S}e^{-\tau\hat{H}_0}] \\ &= \frac{1}{\beta}\int_0^\beta d\tau\,\mathrm{Tr}\left[\frac{e^{-\beta\hat{H}_0}}{Z}\hat{S}\right] = \langle\hat{S}\rangle. \end{aligned} \tag{D2}$$

The susceptibility is then

$$\chi = \frac{\partial M}{\partial h} = \frac{\partial}{\partial h}\left(\frac{1}{\beta}\frac{1}{Z}\frac{\partial Z}{\partial h}\right) = \frac{1}{\beta}\frac{1}{Z}\frac{\partial^2 Z}{\partial h^2} - \frac{1}{\beta}\frac{1}{Z^2}\left(\frac{\partial Z}{\partial h}\right)^2. \tag{D3}$$

The second term is obviously equal to $\beta M^2=\beta\langle\hat{S}\rangle^2$, while the first term can be transformed as follows [37]:

$$\begin{aligned} \frac{1}{\beta}\frac{1}{Z}\left(\frac{\partial^2 Z}{\partial h^2}\right)_0 &= \frac{1}{\beta}\frac{1}{Z}\left(\frac{\partial}{\partial h}\int_0^\beta d\tau\,\mathrm{Tr}[e^{-\beta\hat{H}}\hat{S}]\right)_0 \\ &= \frac{1}{Z}\mathrm{Tr}\left[\left(\frac{\partial e^{-\beta\hat{H}}}{\partial h}\right)_0\hat{S}\right] = \frac{1}{Z}\int_0^\beta d\tau\,\mathrm{Tr}[e^{-\beta\hat{H}_0}\hat{S}(\tau)\hat{S}] \\ &= \int_0^\beta d\tau\langle\hat{S}(\tau)\hat{S}\rangle. \end{aligned} \tag{D4}$$

Thus the susceptibility is given by $\chi=\int_0^\beta d\tau[\langle\hat{S}(\tau)\hat{S}\rangle-\langle\hat{S}\rangle^2]$. From this, the Taylor expansion for Z,

$$Z \simeq Z_0\left\{1+\beta Mh+\frac{1}{2}\beta^2M^2h^2+\frac{1}{2}\beta\chi h^2\right\}, \tag{D5}$$

is identical to the one obtained for the case of mutually commuting Hamiltonians, and therefore a relation analog to Eq. (18) holds between C and the susceptibility χ.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).
[2] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
[3] C. A. Fuchs, Ph.D. thesis, e-print arXiv:quant-ph/9601020.
[4] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, UK, 1999).
[5] P. Zanardi and N. Paunković, Phys. Rev. E **74**, 031123 (2006).
[6] P. Zanardi, M. Cozzini, and P. Giorda, J. Stat. Mech.: Theory Exp. (2007), L02002; M. Cozzini, P. Giorda, and P. Zanardi, Phys. Rev. B **75**, 014439 (2007).
[7] P. Buonsante and A. Vezzani, Phys. Rev. Lett. **98**, 110601 (2007).
[8] H.-Q. Zhou and J. P. Barjaktarevič, e-print arXiv:cond-mat/0701608.
[9] H.-Q. Zhou, J.-H. Zhao, and B. Li, e-print arXiv:0704.2940; H.-Q. Zhou, e-print arXiv:0704.2945.
[10] M. Cozzini, R. Ionicioiu, and P. Zanardi, Phys. Rev. B **76**, 104420 (2007).
[11] A. Hamma, W. Zhang, S. Haas, and D. A. Lidar, e-print arXiv:0705.0026.
[12] S.-J. Gu, H.-M. Kwok, W.-Q. Ning, and H.-Q. Lin, e-print arXiv:0706.2495.
[13] P. Zanardi, P. Giorda, and M. Cozzini, Phys. Rev. Lett. **99**, 100603 (2007).
[14] L. C. Venuti and P. Zanardi, Phys. Rev. Lett. **99**, 095701 (2007).
[15] A. C. M. Carollo and J. K. Pachos, Phys. Rev. Lett. **95**, 157203 (2005); S.-L. Zhu, Phys. Rev. Lett. **96**, 077206 (2006); A. Hamma, e-print arXiv:quant-ph/0602091; M. E. Reuter, M. J. Hartmann, and M. B. Plenio, Proc. R. Soc. London, Ser. A **463**, 1271 (2007).
[16] S. Chen, L. Wang, S.-J. Gu, and Y. Wang, Phys. Rev. E **76**, 061108 (2007).
[17] P. Zanardi, H. T. Quan, X. Wang, and C. P. Sun, Phys. Rev. A **75**, 032109 (2007); P. Zanardi, L. C. Venuti, and P. Giorda, Phys. Rev. A **76**, 062318 (2007).
[18] W.-L. You, Y.-W. Li, and S.-J. Gu, Phys. Rev. E **76**, 022101 (2007).

[19] A. Uhlmann, Rep. Math. Phys. **24**, 229 (1986).
[20] S. Doniach and E. H. Sondheimer, *Green's Functions for Solid State Physicists* (Imperial College Press, London, 1998).
[21] H.-Q. Zhou, R. Orus, and G. Vidal, e-print arXiv:0709.4596.
[22] Argonne National Laboratory Research Library, http://www.library.anl.gov/
[23] P. Cejnar and J. Jolie, Phys. Rev. E **61**, 6237 (2000); P. Cejnar, V. Zelevinsky, and V. V. Sokolov, Phys. Rev. E **63**, 036127 (2001).
[24] M. Hiller, T. Kottos, and T. Geisel, Phys. Rev. A **73**, 061604(R) (2006).
[25] J. Tidström and E. Sjöqvist, Phys. Rev. A **67**, 032110 (2003).
[26] Fidelity is closely related to various distance measures on the set of quantum states: The closer two states are, the more similar (less distinguishable) they are (for an overview of different distance measures induced by the fidelity, see [3]).
[27] The use of the fidelity approach to thermal phase transitions was also suggested in [8].
[28] For the proof of the relation $\varepsilon_F = E_{k_{F\uparrow}} = E_{k_{F\downarrow}}$, see Appendix A.
[29] In the following, we discuss the case of a three-dimensional spatial lattice. This treatment could be easily generalized to a different lattice dimensionality.
[30] Strictly speaking, for $U > U_c$ we speak of multi- (twofold) valued functions $k_{F\uparrow}(U)$ and $k_{F\downarrow}(U)$.
[31] For reasons of simplicity, we use formulas for canonical rather than grand canonical ensemble. Introducing chemical potential results only in a "shift" of the one-particle energies from the unperturbed Hamiltonian $\hat{H}_0$.
[32] Note the similarity, and the difference, to the previously introduced operators $\hat{\Psi}_k$.
[33] We use the well known identity $\cosh(c_k/2) = \sqrt{\frac{1}{2}(1+\cosh c_k)}$.
[34] Alternatively, we could first translate the coordinate system by the vector $(\sqrt{2}A, \sqrt{2}A)$ and then rotate it by the angle $\varphi = \pi/4$ to obtain the same system $(\bar{x}, \bar{y})$.
[35] Using $A^{\pm} = A_x \pm iA_y$ and $A^0 = A_z$, we can write $\vec{A} = A_x\vec{e}_x + A_y\vec{e}_y + A_z\vec{e}_z = A^+\vec{e}_+ + A^-\vec{e}_- + A^0\vec{e}_0$, where $\vec{e}_{\pm} = \vec{e}_x \pm i\vec{e}_y$ and $\vec{e}_0 = \vec{e}_z$ (and analogously for $\vec{B}$). The scalar product between two vectors is then $\vec{A}\cdot\vec{B} = A_xB_x + A_yB_y + A_zB_z = \frac{1}{2}(A^+B^- + A^-B^+) + A^0B^0$. Therefore, $A_k^2 = A^+A^- + (A^0)^2$.
[36] We consider the particular case of the $h=0$ expansion, but all the results obtained can be easily generalized for the $h \neq 0$ case.
[37] In order to avoid a possible confusion, here we explicitly denote that the derivatives are taken for $h=0$.

Ground state overlap and quantum phase transitions

Paolo Zanardi and Nikola Paunković*
Institute for Scientific Interchange (ISI), Villa Gualino, Viale Settimio Severo 65, I-10133 Torino, Italy
(Received 9 January 2006; revised manuscript received 31 May 2006; published 26 September 2006)

We present a characterization of quantum phase transitions in terms of the the overlap function between two ground states obtained for two different values of external parameters. On the examples of the Dicke and *XY* models, we show that the regions of criticality of a system are marked by the extremal points of the overlap and functions closely related to it. Further, we discuss the connections between this approach and the Anderson orthogonality catastrophe as well as with the dynamical study of the Loschmidt echo for critical systems.

DOI: 10.1103/PhysRevE.74.031123 PACS number(s): 64.60.−i, 05.30.−d, 75.10.−b

INTRODUCTION

Quantum phase transitions (QPTs) [1] have drawn a considerable interest within various fields of physics in the recent years. They are studied in condensed matter physics because they provide valuable information about the novel type of finite-temperature states of matter that emerge in the vicinity of QPTs [1]. Unlike the ordinary phase transitions, driven by thermal fluctuations, QPTs occur at zero temperature and are driven by purely quantum fluctuations. In the parameter space, the points of nonanalyticity of the ground state energy density are referred to as critical points and define the QPTs. In these points one typically witnesses the divergence of the length associated to the two-point correlation function of some relevant quantum field. An alternative way of characterizing QPTs is by the vanishing, in the thermodynamical limit, of the energy gap between the ground and the first excited state in the critical points. Recently, a huge interest was raised in the attempt of characterizing QPTs in terms of the notions and tools of quantum information [2]. More specifically QPTs have been studied by analyzing scaling, asymptotic behavior, and extremal points of various entanglement measures [3–7]. More recently, the connection between geometric Berry phases and QPTs in the case of the *XY* model has been also studied [8].

In this paper, we aim to provide yet another characterization of the regions of criticality that define QPTs. We shall show how critical points can be individuated by studying a surprisingly simple quantity: the overlap, i.e., the scalar product, between two ground states corresponding to two slightly different values of the parameters. The physical intuition behind this approach should be obvious: QPTs mark the separation between regions of the parameter space which correspond to ground states having deeply different structural properties, e.g., order parameters. This difference is here quantified by the simplest Hilbert-space geometrical quantity, i.e., the overlap. Note that the square modulus of the overlap is nothing but the fidelity, widely used in quantum information as a function that provides the criterion for distinguishability between quantum states [2]. Therefore it is a natural candidate for a study of macroscopic distinguishability between quantum states that define different macroscopic states of matter (different phases). When applied to cases of many-body systems containing many degrees of freedom, the overlap (or, fidelity) might seem to be too coarse a quantity, and not bearing any apparent information about the difference in order properties between quantum phases to be of any use. Nevertheless, the main result of this paper is that in some cases it is indeed possible to do so. The critical behavior of a system undergoing QPTs is reflected in the geometry of its Hilbert space: approaching the QPTs the overlap (distance) between neighboring ground states shows a dramatic drop (increase). We would like to notice that Cejnar *et al.* [9] discussed the overlap entropy between the eigenbases of a system's Hamiltonian and various physically relevant bases, in the context of enhanced decoherence effects in the regions of criticality (see also Ref. [10]).

In the following two sections, we conduct our analysis on the cases of two simple, yet physically relevant and mathematically instructive, examples of the Dicke model and the *XY* spin-chain model. Next, we discuss the connection between the scaling and asymptotic behaviors and the so-called Anderson orthogonality catastrophe [11]. Moreover, the relation with the dynamical study of decoherence and quantum criticality [12] is briefly addressed. Finally, in the last section conclusions are discussed.

For a generic point in parameter space we use the label $q \in \mathbb{R}^L$, where L is the number of external parameters determining the system's Hamiltonian. As the overlap function depends on the difference between parameters as well, we introduce $\tilde{q} \equiv q + \delta q$ to denote the neighboring point $\tilde{q}$ and the difference δq. Following this notation, we denote the ground states by $|g\rangle \equiv |g(q)\rangle$ and $|\tilde{g}\rangle \equiv |g(\tilde{q})\rangle$. In general, all the functions $F(q)$ evaluated in the point $\tilde{q}$ we will denote as $\tilde{F}$, while those evaluated in the critical point q_c we will denote as F_c [note that by combining two cases, we have $\tilde{F}_c = F(q_c + \delta q)$]. Then, the overlap function is simply given by the scalar product $\langle g(q) | g(\tilde{q}) \rangle$ (note that all the results of this paper could be easily formulated in terms of fidelity as well). We shall examine the behavior of the overlap as a function of q only, while keeping δq fixed and small.

*Present address: SQIG, Instituto de Telecomunicações and Instituto Superior Técnico, P-1049-001, Lisboa, Portugal.

 ©2006 The American Physical Society

DICKE MODEL

Our first example is the Dicke model. It describes a dipole interaction between a single Bosonic mode $\hat{a}$ and a collection of N two-level atoms. If for N atoms we introduce the collective angular momentum operators $\hat{J}_s, s \in \{\pm, z\}$, the Dicke Hamiltonian has the following form (we take $\hbar = 1$):

$$\hat{H}(\lambda) = \omega_0 \hat{J}_z + \omega \hat{a}^\dagger \hat{a} + \frac{\lambda}{\sqrt{2j}}(\hat{a}^\dagger + \hat{a})(\hat{J}_+ + \hat{J}_-). \tag{1}$$

The parameter λ is the atom-field coupling strength and is the one driving the QPTs in this model. Therefore we have $q = \lambda$ and denote the Hamiltonian's dependence on that parameter only. Parameters ω_0 and ω stand for the atomic level-splitting and Bosonic mode frequency, respectively, while j describes the length of a collective spin vector, and is assumed to be constant and equal to $j = N/2$. In the thermodynamical limit ($N \to \infty$), which is here equivalent to ($j \to \infty$), the Dicke Hamiltonian undergoes a quantum phase transition for the critical value of its parameter λ given by $\lambda_c = \sqrt{\omega\omega_0}/2$. When $\lambda < \lambda_c$, the system is in a highly unexcited *normal* phase, while $\lambda > \lambda_c$ defines the *super-radiant* phase in which both the field and N atoms become macroscopically excited. The super-radiant phase is characterized by the broken symmetry given by the parity operator $\hat{\Pi} = \exp(i\pi\hat{N})$, $\hat{N} = (\hat{a}^\dagger \hat{a} + \hat{J}_z + j)$: the ground state is doubly degenerate. As shown in Ref. [13], by introducing Bosonic operators $\hat{b}$ through the Holstein-Primakoff representation [14], the above Dicke Hamiltonian (1) can be exactly diagonalized in the thermodynamical limit. In the normal phase, its form is

$$\hat{H}^n(\lambda) = \omega_0 \hat{b}^\dagger \hat{b} + \omega \hat{a}^\dagger \hat{a} + \lambda(\hat{a}^\dagger + \hat{a})(\hat{b}^\dagger + \hat{b}) - j\omega_0. \tag{2}$$

Its ground state is $g(x,y) = \left(\frac{\varepsilon_+ \varepsilon_-}{\pi^2}\right)^{1/4} e^{-1/2\langle \mathbf{R}, A\mathbf{R}\rangle} [\mathbf{R} = (x,y)]$ where x and y are the real-space coordinates associated to the modes $\hat{a}$ and $\hat{b}$, $A = U^{-1} M U$, $M = \mathrm{diag}[\varepsilon_-, \varepsilon_+]$ and U an orthogonal matrix

$$U = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

[$c = \cos\gamma, s = \sin\gamma$ are given by the squeezing angle $\gamma = (1/2)\arctan[4\lambda\sqrt{\omega\omega_0}/(\omega^2 + \omega_0^2)]$]. $\varepsilon_\pm$ represent the fundamental collective excitations of the system and are given by $\varepsilon_\pm^2 = \frac{1}{2}(\omega^2 + \omega_0^2 \pm \sqrt{(\omega^2 - \omega_0^2)^2 + 16\lambda^2\omega\omega_0})$. From the above formula, we see that $\varepsilon_-(\lambda_c) \equiv \varepsilon_-^c = 0$: the system becomes gapless and undergoes a QPT for $\lambda = \lambda_c$.

The overlap, calculated between two ground states g and $\tilde{g}$, is given by

$$\langle g|\tilde{g}\rangle = 2\frac{[\det A \det \tilde{A}]^{1/4}}{[\det(A + \tilde{A})]^{1/2}} = 2\frac{[\det A]^{1/4}}{[\det \tilde{A}]^{1/4}[\det(1 + \tilde{A}^{-1}A)]^{1/2}}. \tag{3}$$

Note that the overlap is a function of both λ and $\delta\lambda$. In the limit ($\lambda \to \lambda_c$), with $\delta\lambda > 0$ being fixed, $\det A = \varepsilon_+ \varepsilon_- \to 0$, while $\det \tilde{A} \geqslant \det \tilde{A}_c = \tilde{\varepsilon}_+^c \tilde{\varepsilon}_-^c > 0$. The same holds for $\det(1 + \tilde{A}^{-1}A)$, for a sufficiently small $\delta\lambda$ (note that $\lim_{\delta\lambda \to 0} \tilde{A}^{-1} = A^{-1}$). But in the present case, it is possible to show that for every, and not just small, $\delta\lambda$, $\det(1 + \tilde{A}^{-1}A)$ does not vanish. Using the formula $\det(1 + A) = 1 + \mathrm{Tr}A + \det A$ for 2×2 matrices, we get $\det(1 + \tilde{A}^{-1}A) \to 1 + \mathrm{Tr}(\tilde{A}_c^{-1}A_c)$ [note that $\det(\tilde{A}^{-1}A) = [\det \tilde{A}]^{-1} \det A \to 0$]. After a straightforward calculation, we obtain the result: $\mathrm{Tr}(\tilde{A}_c^{-1}A_c) = \frac{\varepsilon_+^c}{\tilde{\varepsilon}_+^c \tilde{\varepsilon}_-^c}[(s\tilde{c} + c\tilde{s})^2 \tilde{\varepsilon}_+^c + (s\tilde{s} - c\tilde{c})^2 \tilde{\varepsilon}_-^c]$. Therefore $\mathrm{Tr}(\tilde{A}_c^{-1}A_c) > 0$ for every λ and we can conclude that $\langle g|\tilde{g}\rangle \propto (\varepsilon_-)^{1/4}$ as ($\lambda \to \lambda_c$). In Ref. [13], it was shown that when approaching the critical point from both normal and super-radiant sides, the excitation energy ε_- drops as the square root of $\Delta \equiv |\lambda - \lambda_c|$, which gives us the asymptotic behavior of the overlap function in the vicinity of the critical point: $\langle g|\tilde{g}\rangle \propto \Delta^{1/8}$. Although we have provided here only the results for overlap function for the system in the normal phase, the analogous analysis for the super-radiant phase gives us the same qualitative results, as the two ground states are again the Gaussian-type states, but with translated and re-scaled x and y axes. Therefore we omit it here.

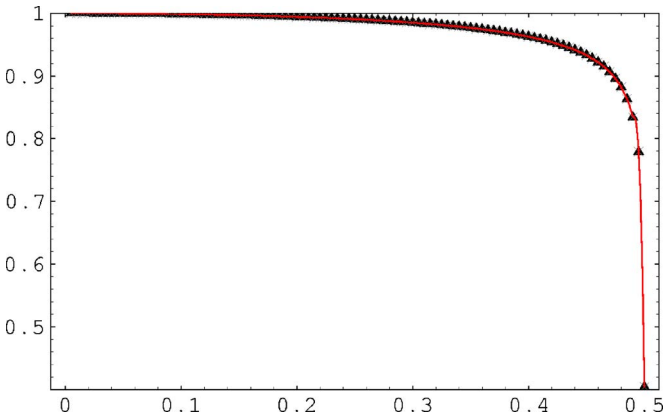


FIG. 1. (Color online) The overlap function $\langle g|\tilde{g}\rangle$, Eq. (3), as a function of $\lambda < \lambda_c$, taken for the resonant case $\omega_0 = \omega = 1$ and $\delta\lambda = 10^{-6}$. Note the dramatic decreasing of the function as we approach the point of criticality.

We conclude this section by presenting the numerical results for the overlap function in the normal phase. In Fig. 1 we plot the overlap (3) between two ground states of the Dicke Hamiltonian for the resonant case $\omega_0 = \omega = 1$ and $\delta\lambda = 10^{-6}$. We see that it is almost constant and equal to unity for a wide range of λ, apart from the very narrow area around λ_c, when it drastically drops to zero. Such behavior of the overlap function around the point of criticality can be ascribed to the fact that the ground state for $\lambda = \lambda_c$ becomes completely delocalized along one of two rotated axes, as opposed to the localized ground state outside of the point of criticality (see Ref. [13]).

XY SPIN CHAIN

In the following section, we discuss the example of the one-dimensional *XY* anisotropic spin-half chain in the exter-

nal magnetic field. Its Hamiltonian is given by the following expression:

$$\hat{H}(\gamma,\lambda) = -\sum_{i=-M}^{M}\left(\frac{1+\gamma}{2}\hat{\sigma}_i^x\hat{\sigma}_{i+1}^x + \frac{1-\gamma}{2}\hat{\sigma}_i^y\hat{\sigma}_{i+1}^y + \frac{\lambda}{2}\hat{\sigma}_i^z\right). \tag{4}$$

The parameter $\gamma \in \mathbb{R}$ defines the anisotropy, while $\lambda \in \mathbb{R}$ represents external magnetic field along the z axis, up to a factor $\frac{1}{2}$. Therefore $q=(\gamma,\lambda)$. The operators $\hat{\sigma}_i^\alpha$, $\alpha \in \{x,y,z\}$ are the usual Pauli operators. This Hamiltonian can be exactly diagonalized by successively applying Jordan-Wigner, Furier, and Bogoliubov transformation (see, for example, Ref. [1]). This way, we obtain the following form of the Hamiltonian: $\hat{H}(\gamma,\lambda)=\Sigma_{k=-M}^{M}\Lambda_k(\hat{b}_k^\dagger\hat{b}_k-1)$. The energies of one-particle excitations are given by $\Lambda_k=\sqrt{\varepsilon_k^2+\gamma^2\sin^2\frac{2\pi k}{N}}$, with $\varepsilon_k=\cos\frac{2\pi k}{N}-\lambda$ and $N=2M+1$ being the total number of sites (spins). One-particle excitations are given by the Fermionic operators $\hat{b}_k=\cos\frac{\theta_k}{2}\hat{d}_k - i\sin\frac{\theta_k}{2}\hat{d}_{-k}^\dagger$, with $\cos\theta_k=\varepsilon_k/\Lambda_k$. Finally, the ground state $|g(\gamma,\lambda)\rangle$, that is defined as the state to be annihilated by each operator $\hat{b}_k$ [$\hat{b}_k|g(\gamma,\lambda)\rangle\equiv 0$], is given as a tensor product of qubitlike states:

$$|g(\gamma,\lambda)\rangle = \bigotimes_{k=1}^{M}\left(\cos\frac{\theta_k}{2}|0\rangle_k|0\rangle_{-k} - i\sin\frac{\theta_k}{2}|1\rangle_k|1\rangle_{-k}\right). \tag{5}$$

In its space of parameters, the family of Hamiltonians given by Eq. (4) exhibits two regions of criticality, defined by the existence of gapless excitations: (i) the XX region of criticality, for $\gamma=0$ and $\lambda\in(-1,1)$; (ii) the XY region of criticality given by the lines $\lambda=\pm 1$.

As in the previous example, let us first consider the exact overlap function. From Eq. (5), it follows that the exact overlap function between the ground states $|g\rangle$ and $|\tilde{q}\rangle$ is

$$\langle g(q)|g(\tilde{q})\rangle = \prod_{k=1}^{M}\cos\frac{\theta_k-\tilde{\theta}_k}{2}, \tag{6}$$

where $\tilde{\theta}_k=\theta_k(\tilde{q})$. Note the dependence on the number of sites N that is implicit in all the previous formulas from this section. In Fig. 2(a), we present the numerical result obtained using the above Eq. (6), for $N=10^6$ spins and $\delta\lambda=\delta\gamma=10^{-6}$. We observe that the regions of criticality are clearly marked by a sudden drop of the value of the overlap function. As before, we ascribe this type of behavior to a dramatic change in the structure of the ground state of the system while undergoing QPTs.

In order to investigate the overlap function more quantitatively and relate its behavior to the existence of the regions of criticality, we note that while the overlap depends on the values of both the parameters q and the difference δq, the regions of criticality are defined by the values of parameters only. Therefore in the following we choose to study the functions

$$S_N^\lambda(\lambda,\gamma) \equiv \sum_{k=1}^{M}\left(\frac{\partial\theta_k}{\partial\lambda}\right)^2, \quad S_N^\gamma(\lambda,\gamma) \equiv \sum_{k=1}^{M}\left(\frac{\partial\theta_k}{\partial\gamma}\right)^2, \tag{7}$$

that define the first nonzero order of the Taylor expansion of the overlap function (6). The functions $S_N^\lambda(\lambda,\gamma)$ and $S_N^\gamma(\lambda,\gamma)$ are natural candidates for our study because they express the "rate of change" of the ground state, taken in the point q. They do not depend on the difference δq, and although for every finite N it is possible to find δq small enough so that the exact overlap is arbitrarily well approximated by the expression $\exp\left[-\frac{1}{8}S_N^q(q)\delta q^2\right]$, the functions $S_N^\lambda(\lambda,\gamma)$ and $S_N^\gamma(\lambda,\gamma)$ on their own capture the behavior of the $\langle g(q)|g(\tilde{q})\rangle$ function and are enough for our current study. They also allow for analytic investigation, together with numerical one. In Figs. 2(b) and 2(c) we present the numerical results for $S_N^\lambda(\lambda,\gamma)$ and $S_N^\gamma(\lambda,\gamma)$, respectively, for $N=10^6$ spins. Again, the regions of criticality could easily be inferred by simply observing both plots. Note that in this case the relative difference between the numerical values in the regions of criticality and elsewhere is much bigger than in the case of the exact overlap [see Fig. 2(a)]. But now, *both* plots are needed to detect both regions of criticality. This is so because by moving along $\gamma=0$, while keeping $|\lambda|<1$, we do not move outside the XX region of criticality and therefore do not expect the qualitative change in the structure of the ground state, and consequently in the behavior of $S_N^\lambda(\lambda,\gamma)$ as well. The same holds for $S_N^\gamma(\lambda,\gamma)$ and the XY region of criticality.

We first examine the scaling behavior of $S_N^\lambda(\lambda,\gamma)$ and $S_N^\gamma(\lambda,\gamma)$ with respect to number of spins N. The numerics present us with the following results. First, as expected, $S_N^\lambda(\lambda,\gamma)$ and $S_N^\gamma(\lambda,\gamma)$ scale linearly with N when $(\gamma\to 0)$ and $(\lambda\to\pm 1)$, respectively. In the regions of criticality, we have that $S_N^\lambda(|\lambda|=1,\gamma)\propto N^2/\gamma^2$ and $S_N^\gamma(|\lambda|\leqslant 1,\gamma=0)\propto N^2$ [we note here that $S_N^\lambda(|\lambda|\leqslant 1,\gamma=0)=0$].

We have also conducted a separate analytical study, confirming the above numerical results. First, we note that for every point q in parameter space, and every *finite* N, partial derivatives $\left(\frac{\partial\theta_k}{\partial\lambda}\right)$ and $\left(\frac{\partial\theta_k}{\partial\gamma}\right)$ are continuous functions of the parameters. They can become infinite only in the thermodynamical limit, when $(N\to\infty)$, and only in the *regions of criticality*. By looking at the explicit form of derivatives (we use $x_k=\frac{2\pi}{N}k$):

$$\left(\frac{\partial\theta_k}{\partial\lambda}\right) = \frac{\gamma(\sin x_k)}{[(\cos x_k-\lambda)^2+\gamma^2(\sin x_k)^2]}, \tag{8}$$

$$\left(\frac{\partial\theta_k}{\partial\gamma}\right) = -\frac{|\sin x_k|(\cos x_k-\lambda)}{[(\cos x_k-\lambda)^2+\gamma^2(\sin x_k)^2]}, \tag{9}$$

we see that only when the energy Λ_k (the denominator of both of the above expressions) gets arbitrarily small (or zero), the derivatives (8) and (9) can become divergent. In other words, only when $\cos x_k$ gets arbitrarily close to λ, *and* either γ or $\sin x_k$ get close to zero. That is, in the regions of criticality. Note that we assume that for every $N\in\mathbb{N}$ and $k\in\{1,\ldots M\}$, equation $\cos x_k=\lambda$ has no solution, which presents a generic case [λ's that allow for the solutions of this

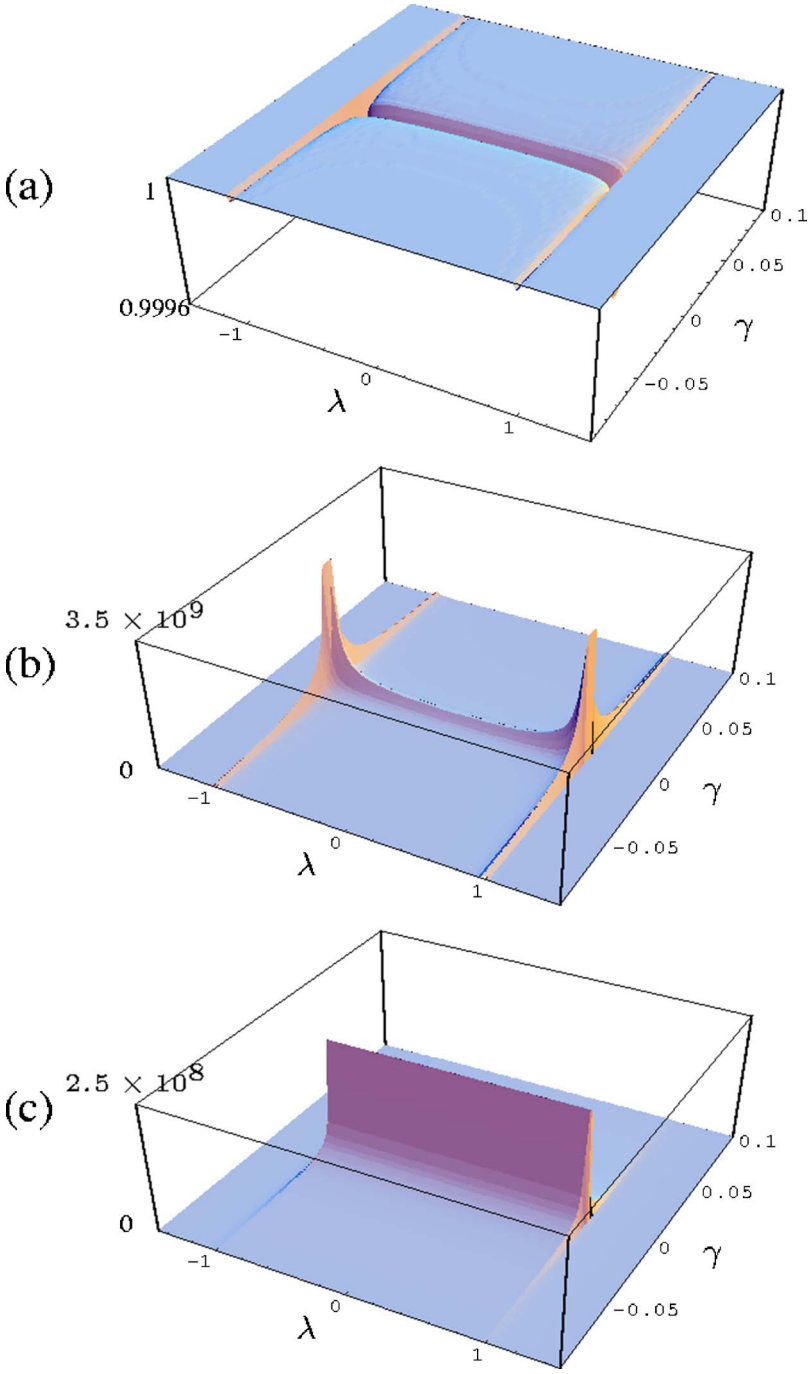


FIG. 2. (Color online) (a) The overlap function $\langle g(q)|g(\tilde{q})\rangle$, as a function of λ and γ, for $N=10^6$ and $\delta\lambda=\delta\gamma=10^{-6}$. Note the clear dips of the plot in the regions of criticality. (b) $S_N^{\lambda}(\lambda,\gamma)$. (c) $S_N^{\gamma}(\lambda,\gamma)$.

equation form a set of measure zero on the $(-1,1)$ interval]. Therefore outside the regions of criticality $S_N^{\lambda}(\lambda,\gamma)$ and $S_N^{\gamma}(\lambda,\gamma)$ scale linearly with N.

Regarding the regions of criticality, we first consider the scaling behavior of $S_N^{\lambda}(\lambda,\gamma)$ in the vicinity of the XX criticality. As there always exists k_0 such that in the $(N\to\infty)$ limit $\cos x_{k_0}\to\lambda$, then for such x_{k_0} and every finite γ, it follows from Eq. (8) that $\left(\frac{\partial\theta_{k_0}}{\partial\lambda}\right)\to(\gamma\sin x_{k_0})^{-1}$, when $(N\to\infty)$. In other words, it does not scale with N [note that although $k_0=k_0(N)$ is a function of N, $\lim_{N\to\infty}\sin x_{k_0}=\sin\arccos\lambda$]. As all other derivatives are finite, we have that $S_N^{\lambda}(|\lambda|<1,\gamma\to 0)\propto N/\gamma^2$. Similar discussion can be applied to the case of $S_N^{\gamma}(\lambda,\gamma)$ in the XY region of criticality. Again, there exists a qubit defined by $k_1=1$ for which $\cos x_{k_1}\to 1$ in the $(N\to\infty)$ limit, so that its existence could bring about the scaling of $S_N^{\gamma}(\lambda,\gamma)$ larger than linear in thermodynamical limit. Using the Taylor expansion of sine and cosine functions around zero (note that in that case, $\sin x_{k_1}\to 0$ as well), from Eq. (9) we obtain $\left(\frac{\partial\theta_k}{\partial\gamma}\right)\propto x_{k_1}/\gamma^2\to 0$, $(N\to\infty)$. In other words, $S_N^{\gamma}(|\lambda|=1,\gamma)\propto N$.

Now, we turn to more interesting cases of the relevant functions $S_N^{\lambda}(\lambda,\gamma)$ and $S_N^{\gamma}(\lambda,\gamma)$, in the XY and XX regions of criticality, respectively. Using Taylor expansions of sine and cosine functions around zero, we see that in $\lambda=\pm 1$ the derivative $\left(\frac{\partial\theta_{k_1}}{\partial\lambda}\right)$ given by $k_1=1$ behaves like $\left(\frac{\partial\theta_{k_1}}{\partial\lambda}\right)\propto N/(2\pi\gamma)$ as $(N\to\infty)$ [see Eq. (8)] and therefore $S_N^{\lambda}(|\lambda|=1,\gamma)\propto N^2/\gamma^2$. We

also see that the scaling factor depends on γ. Finally, from Eq. (9), we also see that $S_N^{\gamma}(|\lambda| \leqslant 1, \gamma=0) \propto N^2$.

The alternative way to examine the signatures of QPT is to look at the asymptotic behavior of two functions (7) near the regions of criticality. From the numerical study we obtain that the asymptotic behavior of $S_N^{\lambda}(\lambda, \gamma)$ in the vicinity of critical points $\lambda_c = \pm 1$, $\gamma \in (0,1]$, is given by the following formula: $S_N^{\lambda}(\lambda, \gamma) \propto a(\gamma, N)/|1-\lambda|^{\alpha(\gamma,N)}$. From the study of the scaling behavior, we already know that $a(\gamma, N) = a(\gamma) N^2$ and that $a(\gamma) \propto 1/\gamma^2$. Further, from numerics we have that the exponent $\alpha(\gamma, N)$ is constant with respect to γ and approaches to $\alpha = 1$ as $(N \to \infty)$. Such asymptotic behavior, with constant exponent $\alpha = 1$ for all $\gamma \in (0,1]$ could be seen as a consequence of the fact that in that range of parameters the *XY* model belongs to the same class of universality. The numerics gives also the asymptotic behavior of $S_N^{\gamma}(\lambda, \gamma)$ in the vicinity of $\gamma = 0$ (with $|\lambda| < 1$) similar to the previous one, $S_N^{\gamma}(\lambda, \gamma) \propto b(\lambda, N)/\gamma^{\beta(\lambda,N)}$, with the exponent $\beta(\lambda, N)$ approaching to $\beta = 1$ as $(N \to \infty)$. But, the coefficient $b(\lambda, N)$ depends only on N, and as noted before, scales linearly with it, $b(\lambda, N) \propto bN$.

QPT: ORTHOGONALITY CATASTROPHE, LOSCHMIDT ECHO

The above two examples represent a generic case of a many-body system which exhibits continuous QPTs only in the thermodynamical limit. In the case of the *XY* model, as the number of spins increases, the overlap between two different ground states (5) approaches zero, no matter how small the difference in parameters δq is, so that in thermodynamical limit each two ground states are mutually orthogonal; they live in a continuous tensor product space [15]. Such behavior of systems having infinitely many degrees of freedom, when the two physical states corresponding to two arbitrarily close sets of parameters (two arbitrarily similar physical situations) become orthogonal to each other, has been already studied in many-body physics and is known as Anderson *orthogonality catastrophe* [11]. From our study of the *XY* model, we have seen not only that every two ground states become orthogonal in thermodynamical limit, but also the rate of "orthogonalization" between two ground states of large, but finite system, changes qualitatively and grows faster in the vicinity of the regions of criticality. This way, the regions of criticality of an infinite system are already marked by the scaling and asymptotic behavior of the relevant functions of a finite-size system. Loosely speaking, the regions of criticality of QPTs are given as regions where the orthogonality catastrophe is expressed on a qualitatively greater scale. Notice that recently, the occurrence of a particular instance of Anderson-type orthogonality catastrophe was studied for the case of a system in the vicinity of QPTs [16].

Now we would like to establish an explicit connection between the sort of kinematical approach used in this paper and the dynamical one of Ref. [12]. In order to do so let us introduce the projected density of states function $D(\omega; q, \tilde{q}) \equiv \langle g(\tilde{q}) | \delta[\omega - \hat{H}(q)] | g(\tilde{q}) \rangle$ that describes the spread of the ground state $|g(\tilde{q})\rangle$ expressed in the eigenbasis obtained for the point q. Then, the square of the overlap can be expressed as

$$|\langle g(q)|g(\tilde{q})\rangle|^2 = 1 - \int_{E_1}^{\infty} D(\omega) d\omega \tag{10}$$

(E_1 denotes the first excited eigenvalue). We see that in regions in which the spread of $|g(\tilde{q})\rangle$ with respect to $|g(q)\rangle$ is big [quantified in terms of the variance of $D(\omega)$], in other words, in regions where two ground states differ significantly, the overlap will be small. Recently, Quan *et al.* [12] established a link between the critical behavior of the environment and quantum decoherence, showing that the Loschmidt echo [17,18] $L(q,t)$ of the environment exponentially goes to zero as we approach the regions of criticality. A simple algebra gives us that the Fourier transform of the projected density of states is precisely the Loschmidt echo, $\left|\int_{-\infty}^{+\infty} D(\omega) e^{-i\omega t} d\omega\right|^2 = L(q,t)$. We see that the kinematics of a system, given by the geometry of ground states through the overlap function, influences its dynamics as well: the smaller the overlap, i.e., broader $D(\omega)$, the faster the decay of the Loschmidt echo.

CONCLUSIONS AND DISCUSSION

In this paper, by discussing the examples of the Dicke and *XY* spin-chain models, we have presented a characterization of quantum phase transitions based on the study of the scaling and asymptotic behaviors of the overlap between two ground states taken in two close points of the parameter space. Though this quantity might in general not provide an efficient numerical tool it is conceptually quite appealing. In fact the ground state overlap is a purely Hilbert-space geometrical quantity, whose investigation does not rely on any *a priori* understanding of the specific kind of order patterns or peculiar dynamical correlations hidden in the analyzed system. While in the case of the Dicke Hamiltonian it was possible to analyze the overlap function directly in the thermodynamical limit, the case of the *XY* model is more subtle. The process of "orthogonalization" between two ground states has two physically different mechanisms in the case when two states belong to two different phases: one is a common decrease of the overlap due to infinite number of subsystems in thermodynamical limit, the other is characteristic for the case of QPTs and is due to different structures of ground states in different quantum phases.

ACKNOWLEDGMENTS

N.P. was funded by the European Commission, Contract No. IST-2001-39215 TOPQIP. The authors greatly acknowledge the discussions with A. T. Rezakhani.

[1] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, England, 1999).
[2] For reviews, see A. Steane, Rep. Prog. Phys. **61**, 117 (1998); D. P. DiVincenzo and C. H. Bennett, Nature (London) **404**, 247 (2000).
[3] A. Osterloh, L. Amico, G. Falci, and R. Fazio, Nature (London) **416**, 608 (2002).
[4] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003).
[5] Y. Chen, P. Zanardi, Z. D. Wang, and F. C. Zhang, New J. Phys. **8**, 97 (2006).
[6] S.-J Gu, G.-S. Tian, and H.-Q. Lin, quant-ph/0509070.
[7] L.-A. Wu, M. S. Sarandy, and D. A. Lidar, Phys. Rev. Lett. **93**, 250404 (2004).
[8] A. C. M. Carollo and J. K. Pachos, Phys. Rev. Lett. **95**, 157203 (2005); Shi-Liang Zhu, *ibid.* **96**, 077206 (2006); A. Hamma, quant-ph/0602091.
[9] P. Cejnar, V. Zelevinsky, and V. V. Sokolov, Phys. Rev. E **63**, 036127 (2001).
[10] A. Volya and V. Zelevinsky, Phys. Lett. B **574**, 27 (2003); V. Zelevinsky, B. A. Brown, N. Fraizer, and M. Horoi, Phys. Rep. **276**, 85 (1996); P. Cejnar and J. Jolie, Phys. Rev. E **58**, 387 (1998); **61**, 6237 (2000).
[11] P. W. Anderson, Phys. Rev. Lett. **18**, 1049 (1967).
[12] H. T. Quan, Z. Song, X. F. Liu, P. Zanardi, and C. P. Sun, Phys. Rev. Lett. **96**, 140604 (2006).
[13] C. Emary and T. Brandes, Phys. Rev. E **67**, 066203 (2003).
[14] T. Holstein and H. Primakoff, Phys. Rev. **58**, 1098 (1949).
[15] J. von Neumann, Comput. Appl. Math. **6**, 1 (1938); T. Thiemann and O. Winkler, Class. Quantum Grav. **18**, 4997 (2001).
[16] S. Sachdev, M. Troyer, and M. Vojta, Phys. Rev. Lett. **86**, 2617 (2001).
[17] P. R. Levstein, G. Usaj, and H. M. Pastawski, J. Chem. Phys. **108**, 2718 (1998); G. Usaj, H. M. Pastawski, and P. R. Levstein, Mol. Phys. **95**, 1229 (1998); H. M. Pastawski, P. R. Levstein, G. Usaj, J. Raya, and J. Hirschinger, Physica A **283**, 166 (2000).
[18] Z. P. Karkuszewski, C. Jarzynski, and W. H. Zurek, Phys. Rev. Lett. **89**, 170405 (2002); F. M. Cucchietti, D. A. R. Dalvit, J. P. Paz, and W. H. Zurek, *ibid.* **91**, 210403 (2003).

Operational verification of the existence of a spacetime manifold

Nikola Paunković[1] and Marko Vojinović[2]

[1]Instituto de Telecomunicaĉoes and Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001, Lisboa, Portugal

[2]Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

arXiv:2209.04783v3 [gr-qc] 20 May 2025

We argue that there exists an operational way to establish the observability of the notions of space and time. Specifically, we propose a theory-independent protocol for a gedanken-experiment, whose outcome is a signal establishing the observability of the spacetime manifold, without a priori assuming its existence. The experimental signal contains the information about the dimension and the topology of spacetime (with the currently achievable precision), and establishes its manifold structure, while respecting its underlying diffeomorphism symmetry. We also introduce and discuss appropriate criteria for the concept of emergence of spacetime, which any tentative theoretical model of physics must satisfy in order to claim that spacetime does emerge from some more fundamental concepts.

1 Introduction

Both in physics and philosophy, one often asks deep fundamental questions such as "What is time?", "What is space?", "Do they objectively exist or not?", and similar. While a lot has been said and written about these questions throughout history of science and philosophy [1, 2, 3, 4, 5, 6, 7, 8], there appear to be very few attempts to address these questions in terms of *experimental evidence* for the notions of time and space. On one hand, we all have intuitive feeling for both, since we heavily rely on them in daily life, and this intuition is partially based on some experimental evidence. On the other hand, conceptual analysis of the *objective existence* of space and time turned out to be quite a hard problem, due to the high level of symmetry properties of space and time, encoded in the principle of general relativity.

Namely, there is a long-standing common claim that the points on a spacetime manifold are physically unobservable, given the principle of general relativity, i.e., because we expect all physics to be diffeomorphism-invariant and background-independent. As far as the statement goes, one cannot distingush between "this point" versus "that point" of spacetime itself, but only "the point where fields have this value" versus "the point where fields have that value", along the lines of the paradigm of relational approach to physics (see for example [9]).

Loosely speaking, the basic argument for the unobservability of individual spacetime points goes as follows. If we choose one point of spacetime (by specifying its coordinates in some coordinate system), determine the values of all fields at that point, and then perform an "active diffeomorphism" (permutation of manifold points), we "move" all physics from that point to another point. After that, we can perform a "passive diffeomorphism" (choice of a different manifold chart), to undo the active one, i.e., we use the same set of numbers as coordinates for the new point in the new coordinate chart as we have used for the old point in old coordinates. Given that physics does not change throughout the process, we conclude that one cannot distinguish between the "old spacetime point" and the "new spacetime point". Thus, individual spacetime points are unobservable.

While this is correct in itself, apparently there are proposals that go even further, and generalize this argument to deny the existence of the spacetime manifold altogether. The argument could possibly be paraphrased in the following form — if spacetime points are not observable, they do not objectively exist, so therefore the manifold itself does not objectively exist.

The purpose of this work is to challenge this generalization. Namely, our statement is the following: the fact that we cannot observe individual points does *not* imply that the spacetime manifold as a whole cannot be observed, or that it does not objectively exist in nature. In particular, a manifold has properties which are invariant with respect to diffeomorphisms — specifically, its dimension and its topology, and a priori both of these might in principle be observable. The main point of this work is to demonstrate that these properties of a manifold *indeed are observable*. As a consequence of this, we argue that the *spacetime manifold is itself observable*, without violating either the diffeomorphism symmetry or background independence. Additionally, our argument for the observability of spacetime, presented below, relies on an operational, theory-independent experimental pro-

Nikola Paunković: npaunkov@math.tecnico.ulisboa.pt
Marko Vojinović: vmarko@ipb.ac.rs

tocol, incorporated in a specific proposed gedanken-experiment. It is thus mainly based on experimental evidence, rather than on some kind of theoretical or metaphysical assumptions.

In order to further emphasise this last point, one can imagine a hypothetical scenario involving an artificial inteligence (AI) implemented within a memory of a computer, without having any a priori notion about time and space. One can further imagine that this AI is perfectly capable of performing complex mathematical analyses. In such a setup, the experimenter could perform our proposed protocol, and feed AI the resulting experimental data to analyse it. The outcome of the analysis performed by the AI would then be a conclusion that the experimenter "lives in a space and time" of dimension 4 and simply connected topology, despite the fact that AI itself does not have any intuitive or a priori notions of either space or time. This hypothetical scenario emphasises the independence of our proposed experimental protocol on any a priori notions of time and space that an expermenter may be subject to, because the AI would reach the same conclusion without such a priori notions.

The layout of the paper is as follows. Sections 2 and 3 discuss the details of the thought experiment in mechanics and field theory, respectively. The two cases have been separated into two sections merely for the pedagogical purpose of gradually introducing the relevant analysis and techniques, while in principle the analysis in the context of mechanics is merely a special case of the analysis in the context of field theory. Section 4 contains a discussion of various related aspects of the analysis, as well as a discussion of several topics which put our results in a wider context. Among other topics, we discuss the notion of spacetime emergence. In particular, we provide non-trivial criteria that arguably have not been satisfied by any of the existing proposals aiming to describe the emergence of spacetime. Appendix A contains some important technical results needed for the analysis.

2 Mechanics and time

In order to have a clear understanding of the ideas proposed in this work, it is prudent to first discuss the toy-example of a $(0+1)$-dimensional spacetime manifold, i.e., the time manifold. Various key properties of the analysis can be discussed using simple mechanical systems, such as pendulums, while the proper $(3+1)$-dimensional spacetime manifold is postponed for Section 3.

Of course, throughout the text we uphold an initial assumption that the concepts of time and spacetime are not observable a priori.

2.1 System with one observable

Let us introduce the following gedanken-experiment. We are given a swinging pendulum, denoted A. Without any assumptions about Newtonian mechanics (since it relies on the a priori notion of time), our objective is to describe the motion of the pendulum, as precisely as we can. The measurement apparatus we have available is a camera that can take still photos of the pendulum, along with a ruler that can measure the distances on the photos.

Start by taking N photos of the pendulum, completely randomly, and use a ruler to measure the signed distance a_k of the pendulum from its vertical axis, for every photograph $k \in \{1, \dots, N\}$. The distance is signed in the sense that if the pendulum is left of its axis we consider the distance to be negative, while if the pendulum is right of its axis, the sign of distance is positive (the choices of left and right are merely a convention and do not influence any conclusions). Moreover, we want to eliminate any information about the "time order" in which the photos might have been taken, so we remember only the following *unordered* set of measurements, describing what we can tell about the motion of pendulum A,

$$A = \{a_1, a_2, \dots, a_N\}\,, \tag{1}$$

where

$$a_k \in S_A \equiv [a_{min}, a_{max}] \subset \mathbb{R}\,, \quad \forall k = 1, \dots, N\,. \tag{2}$$

Here a_{min} and a_{max} denote the left and right amplitudes of the pendulum, respectively. In usual circumstances one should obtain $a_{min} = -a_{max}$, but for our purposes this equality does not really matter. Also, we assume that no two measurements produce exactly equal values of a, so that $a_i \neq a_j$ for all measurements i and j. This assumption is just for convenience, and we will discuss later the conceptual case in which the measurement results are repeating themselves, and always falling into a discrete set of outcomes.

Given the measurement dataset A, one can draw it as a one-dimensional scatter plot on the a-axis, which looks similar to the plot in Figure 1. If we take additional N measurements (i.e., a total of $2N$), the scatter plot will look similar, only with twice as many points in the segment S_A. In the limit $N \to \infty$, the scatter plot will become dense, and practically fill the whole segment S_A, given that we can measure distances a_k with an arbitrarily small but still nonzero precision. The notion of arbitrary small distance is not in conflict with quantum mechanics, since we are measuring only the position of the pendulum, and

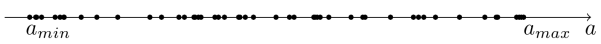


Figure 1: A scatter plot for a single pendulum system.

not its momentum. With sufficiently many measurements, we can describe the motion of the pendulum using the whole set S_A of possible positions it might be in, and we can call this set the *configuration space* of the pendulum.

However, our naive intuition suggests that such a description of pendulum's motion is not completely satisfactory, since each photo demonstrates that the pendulum is at some particular distance a, while the configuration space of the pendulum describes the pendulum merely "in all positions", suggesting that we are missing some extra information. We therefore ask for a more precise description of the pendulum's motion. The precise formalization of this "intuition" will be given below, but at this stage let us appeal to the ideas of relationalism, and try to obtain the missing information by *comparing* the motion of the pendulum against *some other physical system.*

2.2 System with two observables

For lack of a better idea, let us introduce another pendulum, denoted B, in addition to A. In a generic situation, the pendulum B may have different length and other properties compared to A (we discuss this and other properties of the protocol setup in more detail in Subsection 2.3). We put the two pendulums side by side, and randomly take photos of the whole system. We then cut each photo into two pieces, such that each pendulum is displayed separately on its half of the photo. As before, we now shuffle all photos to erase the order in which the photos were taken, as well as the relation which photo of A is paired to which photo of B. We end up with the following *unordered* datasets for the two pendulums:

$$A = \{a_1, a_2, \ldots, a_N\}\,, \quad B = \{b_1, b_2, \ldots, b_N\}\,. \tag{3}$$

After sufficiently many measurements, we can establish the individual configuration spaces for each pendulum like before,

$$S_A \equiv [a_{min}, a_{max}] \subset \mathbb{R}\,, \quad S_B \equiv [b_{min}, b_{max}] \subset \mathbb{R}\,. \tag{4}$$

Again, we have an intuition that this is not a complete description of the motion of the two pendulums. But now we can formalize this intuition as follows. Given the individual configuration spaces S_A and S_B, a priori one can say that the *joint* configuration space will be the set $S_A \times S_B$. One can visualize this experimental result within this set graphically, as follows. Pick a random permutation π of N elements, and construct the following $2 \times N$ matrix,

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_N \\ b_{\pi(1)} & b_{\pi(2)} & \ldots & b_{\pi(N)} \end{pmatrix}, \tag{5}$$

which describes an arbitrary "pairing" of each measurement a_k to some measurement $b_{\pi(k)}$. These pairings define N points on a scatter plot in the space $S_A \times S_B$, which will typically look like the one in Figure 2. One can draw a similar scatter plot for every choice of the permutation function π. Since for N measurements there are $N!$ possible permutations, there are also $N!$ scatter plots. But, as an experimental result of our thought experiment, among all those scatter plots, there will be *one special plot*, corresponding to the permutation denoted as $\tilde{\pi}$, which is visually very distinguishable from the rest, and looks like the one in Figure 3. Looking at the plot, it is straightforward to interpret it as a set of points on a 1-dimensional curve, denoted $\mathcal{C}$. The experimental fact that the joint configuration space $S_{A\cup B}$ is equal to the curve $\mathcal{C}$ has one important consequence — it captures our intuition of a more precise description of the motion of pendulums A and B. This is because the curve $\mathcal{C}$ is equivalent to a functional dependence

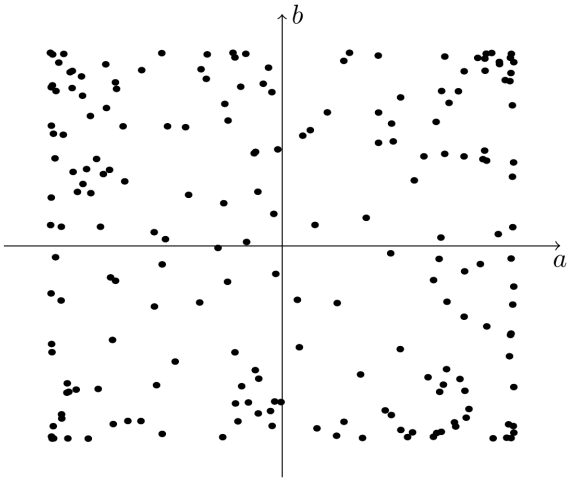


Figure 2: An ordinary scatter plot for a double pendulum system.

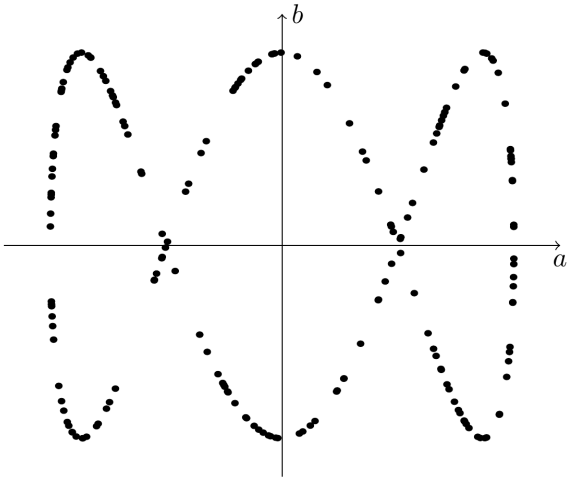


Figure 3: An extraordinary scatter plot for a double pendulum system.

(a *correlation*) between observables a and b, of the general form

$$F(a,b) = 0\,. \tag{6}$$

The equation (6) actually defines the curve $\mathcal{C}$ as a set of all pairs $(a,b) \in S_A \times S_B$ which satisfy (6). At this point we can spell out the formalization of our intuition for the results of the thought experiment, in the form of the following conjecture.

Conjecture 1 *(definition of special correlations). The joint configuration space $S_{A\cup B}$ is a strict subspace of $S_A \times S_B$, which is of measure zero compared to $S_A \times S_B$:*

$$S_{A\cup B} \subset S_A \times S_B\,, \qquad \frac{\mathrm{meas}(S_{A\cup B})}{\mathrm{meas}(S_A \times S_B)} = 0\,. \tag{7}$$

The measure over these sets is defined as an induced measure from the standard Euclidean metric over $\mathbb{R}^2$.

Note that, while visual distinguishability is appealing, it is a particular pattern-recognition trait of a human brain. In order to have a more formal definition of why this particular diagram is "special", without appealing to our eyes and brains, we can resort to the statistical analysis described in detail in the Appendix A. This can be implemented as a computer algorithm operating on dataset (3), eliminating any reliance on visual inspection of scatter plots. Also, the formal statistical analysis has an advantage of being applicable to higher dimensions, i.e., beyond the 2- and 3-dimensional scatter plots, in contrast to visual inspection by a human. This property will become important in Section 3.

The "special" permutation $\tilde{\pi}$, corresponding to the "special" plot above, has several crucial properties:

Conjecture 2 *(properties of the special correlations).*

- Existence. *Given a completely random set of numbers a_k and b_k, $k = 1,\ldots,N$, the permutation $\tilde{\pi}$ does not necessarily exist. This is actually one of the definitional properties for a sequence of numbers to be "random" — absence of any noticable correlation. It should be stressed that the existence of $\tilde{\pi}$ is a specific property of the dataset (3), which* fails to be random*, as opposed to an arbitrary set of numbers not obtained by experimental measurements of the two pednulums. The existence of $\tilde{\pi}$ is therefore* an experimental signal.

- Self-reinforcement. *Suppose that, after taking N measurements (3), we continue to take additional M measurements,*

$$\begin{aligned} A' &= \{a_{N+1}, a_{N+2}, \ldots, a_{N+M}\}\,, \\ B' &= \{b_{N+1}, b_{N+2}, \ldots, b_{N+M}\}\,. \end{aligned} \tag{8}$$

  We can then perform the same analysis independently over the three datasets — the original dataset (A,B), the new dataset (A',B'), and their union dataset $(A\cup A', B\cup B')$, to arrive at the three "special" permutations $\tilde{\pi}$, $\tilde{\pi}'$ and $\tilde{\Pi}$, defined over N, M and $N+M$ numbers respectively. Then, it is again an experimental result *that the restriction of $\tilde{\Pi}$ to the first N numbers will be equal to the original permutation $\tilde{\pi}$, while the restriction of $\tilde{\Pi}$ to the remaining M numbers will be equal to the new permutation $\tilde{\pi}'$. Thus, we have a piecewise-defined equation*

$$\tilde{\Pi}(k) = \begin{cases} \tilde{\pi}(k)\,, & k = 1,\ldots,N\,, \\ \tilde{\pi}'(k)\,, & k = N+1,\ldots,N+M\,, \end{cases} \tag{9}$$

  which holds for all possible choices of N and M. Pictorially, this property means that once we draw scatter plots corresponding to $\tilde{\pi}$ and $\tilde{\pi}'$, the data points will remain nicely aligned along the same curve*. In other words, the correlation which represents our experimental signal* reinforces itself *when one adds additional data.*

  Note that self-reinforcement is an important property in situations where one is interested in discussing the discrete-to-continuum limit. In the context of spacetime, the discrete-to-continuum limit has been extensively studied in causal set theory. See for example [10] for details.

- Dimensionality. *In the limit $N \to \infty$, one can see that there exists a special permutation $\tilde{\pi}$ such that the data points are arranged in a dense way along the whole curve $\mathcal{C}$. In other words, the actual* joint configuration space *$\mathcal{C} \equiv S_{A\cup B}$ for pendulums A and B is not only a subset of measure zero in the direct product $S_A \times S_B$ of two individual configuration spaces, but also a* 1-dimensional subset*. This is also an experimental result. Namely, it could have happened that $S_{A\cup B}$ is some patch in $S_A \times S_B$ of nonzero area (2-dimensional), or a discrete set of points (0-dimensional), or any combination thereof. But these alternative scenarios may happen only for non-experimental datasets, while for experimental datasets it always turns out that $S_{A\cup B}$ is* 1-dimensional.

- Topology. *The statistical analysis from Appendix A provides us not only with the dimension, but also with a convenient set of patches along the curve $\mathcal{C}$, which define a* basis of open sets*, giving rise to an induced* topology *on $\mathcal{C}$. In the example of the curve in the picture, the topology is that of a circle, but it can also be open curve in other examples, as we shall discuss in more detail below.*

Once the existence and properties of the 1-dimensional curve $\mathcal{C}$ have been established, one may be tempted to call it the "time manifold". However, there are several issues associated with that, discussed in the next Subsection.

2.3 Interpreting the results as a time manifold

At the operational level, our protocol makes use of a device which takes a photo of two pendulums. One immediate question that arises regarding this setup is whether this implicitly assumes some a priori notion of "simultaneity", or maybe even "time". Namely, taking a single photo of two pendulums may be considered equivalent to "simultaneously" taking two photos of two pendulums individually. However, this is not equivalent — we deliberately take *one single* photo of a *composite system*, rather than *two separate* photos of its *subsystems*. This is because the requirement that the two separate photos be taken "simultaneously" would indeed necessarily assume an a priori notion of time. On the other hand, one can circumvent any such assumption by taking a single photo to *operationally define* the notion of "simultaneity", without a priori assuming time. Our definition is as follows: we use a camera to produce a single photo of a measured system — as long as one can decompose such a system into two parts, one is performing a *simultaneous measurement* of those two parts. With respect to this definition, it is immaterial whether these two parts are two independent pendulums, or maybe two pieces of a single pendulum, or otherwise. By taking a single photo of the whole system, any two of its subsystems are said to have been photographed "simultaneously".

Another question one can ask regarding our experimental setup is the following. In our protocol, after taking a single photo of two pendulums we cut it into two pieces, one for each pendulum. Alternatively, we could have discussed a different protocol, in which we would simply take two photos, one for each pendulum, instead of taking one photo and cutting it in two. As it turns out, if one collects the data using this alternative protocol, one will fail to detect the extraordinary scatter plot (from Figure 3) when analysing the data. Taking a single photo and cutting it in two parts is thus a crucial feature of our protocol, one that enables us to detect the relevant signal. That is why we have deliberately chosen to introduce the protocol as was done in the previous Subsection.

A further issue that needs to be discussed is the question whether it is possible to obtain the desired signal in the case when the two pendulums are mutually interacting in some way. In order to address this question, let us consider an example, a situation with two electrically charged pendulums, so that they interact electromagnetically while swinging. Repeating the procedure of taking photos and analysing, like before, the new datasets A and B, one obtains the results which differ from the ones above only in the details of the shape of the curve $\mathcal{C}$. This is because the interaction between the two pendulums will influence their relative positions and motion, but the extraordinary permutation similar to the one in Figure 3 will still exist, and the curve obtained from it will qualitatively remain the same, namely a 1-dimensional curve. This is a generic result, independent of the type and details of the interaction between pendulums.

Next, in our case of the two pendulums the curve $\mathcal{C}$ is almost certainly self-intersecting, preventing it from being a manifold in the proper sense. Second, if we denote the length of the pendulums A and B as l_A and l_B respectively, it may happen that their square-roots are incommensurate,

$$\sqrt{\frac{l_A}{l_B}} \notin \mathbb{Q}\,. \tag{10}$$

Note that if we recall Newtonian mechanics, (10) in fact means that the periods of the two pendulums are incommensurate. Of course, in order to be consistent, we have to assume that we have no knowledge about either Newtonian mechanics nor about the concept of a "period", so we refrain from this interpretation. Nevertheless, we are still allowed to measure the lengths l_A and l_B using a ruler, calculate their square roots, and discuss the validity of (10) in a given experiment. And if it happens that (10) holds, in the limit $N \to \infty$ the curve $\mathcal{C}$ becomes a space-filling curve (passing through almost every point in the rectangle $S_A \times S_B$), invalidating the statement that it is 1-dimensional, in the sense of the analysis discussed in Appendix A. In order to work around these two issues, it is useful to enlarge our physical system yet again.

2.4 Systems with three and more observables

In addition to pendulums A and B, let us introduce yet another pendulum, denoted C, and repeat the whole analysis, taking photos of all three pendulums side-by-side. This time we end up with the dataset

$$\begin{aligned} A &= \{a_1, a_2, \ldots, a_N\}\,, \\ B &= \{b_1, b_2, \ldots, b_N\}\,, \\ C &= \{c_1, c_2, \ldots, c_N\}\,. \end{aligned} \tag{11}$$

After sufficiently many measurements, we can establish the individual configuration spaces for each pendulum, as before,

$$\begin{aligned} S_A &\equiv [a_{min}, a_{max}] \subset \mathbb{R}\,, \\ S_B &\equiv [b_{min}, b_{max}] \subset \mathbb{R}\,, \\ S_C &\equiv [c_{min}, c_{max}] \subset \mathbb{R}\,. \end{aligned} \tag{12}$$

We can now expect that the joint configuration space will be a measure-zero subset of the 3-dimensional box $S_A \times S_B \times S_C$. To see this, introduce two arbitrary permutations of N elements, π and ρ, and construct the following $3 \times N$ matrix,

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_N \\ b_{\pi(1)} & b_{\pi(2)} & \ldots & b_{\pi(N)} \\ c_{\rho(1)} & c_{\rho(2)} & \ldots & c_{\rho(N)} \end{pmatrix}, \tag{13}$$

which describes an arbitrary ordering of our dataset into triplets $(a_k, b_{\pi(k)}, c_{\rho(k)})$, $k = 1, \ldots, N$. We use these triplets as data points in the 3-dimensional scatter plot whose domain is the box $S_A \times S_B \times S_C$. One can construct $(N!)^2$ such scatter plots, one for each choice of the permutations π and ρ. A typical plot will have data points randomly distributed and filling up the whole box, as in Figure 4. However, either by using visual inspection over $(N!)^2$ such diagrams, or by using the data-analysis approach from Appendix A, one can again establish the existence of the "special" diagram, see Figure 5, corresponding to the "special" permutations $\tilde{\pi}$ and $\tilde{\rho}$.

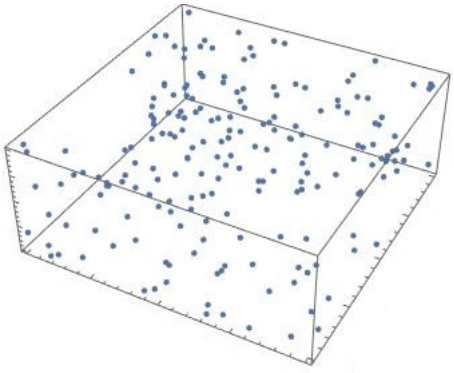


Figure 4: An ordinary scatter plot for a triple pendulum system.

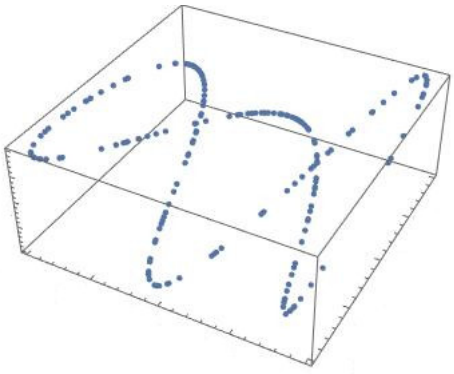


Figure 5: An extraordinary scatter plot for a triple pendulum system.

There are now several new features to be discussed. To begin with, as an experimental result of our thought-experiment, we can establish that our dataset (A, B, C) together with the permutations $\tilde{\pi}$ and $\tilde{\rho}$, satisfies all three previously discussed properties of *existence*, *self-reinforcement* and *dimensionality*. Therefore, the dataset again gives rise to a 1-dimensional curve $\mathcal{C}$, as an *an experimental result*. In principle, the addition of the third pendulum could have rendered the data-points distributed along some 2-dimensional surface in the box rather than a 1-dimensional curve, while still being compatible with the 1-dimensional curve if one only looks at the projection defined as the (A, B) subset of the data. But such a scenario did not happen, and instead the data points are still aligned along a 1-dimensional curve $\mathcal{C}$. Therefore, it is a genuine experimental signal that the joint configuration space has 1-dimensional structure. Moreover, this conclusion survives if one enlarges the physical system even further, by adding additional pendulums, or any other (even aperiodic) mechanical systems. From the point of view of our thought experiment, it is a completely general result.

Second, the 1-dimensional nature of the curve $\mathcal{C}$ is equivalent to the set of two correlation functions between three observables a, b, c,

$$F_1(a, b, c) = 0\,, \qquad F_2(a, b, c) = 0\,. \tag{14}$$

In general, if we have a very big mechanical system of K observables $a^{(1)}, \ldots, a^{(K)}$, we will find in total $K - 1$ correlation functions

$$\begin{aligned} F_1(a^{(1)}, \ldots, a^{(K)}) &= 0\,, \\ &\vdots \\ F_{K-1}(a^{(1)}, \ldots, a^{(K)}) &= 0\,, \end{aligned} \tag{15}$$

whose set of solutions describes a 1-dimensional curve $\mathcal{C}$ in a K-dimensional configuration space.

Next, given the curve $\mathcal{C}$, one can project it onto the (a, b) plane, simply by ignoring the value of the observable c, so that one recovers a curve defined by the points $(a_k, b_{\tilde{\pi}(k)})$ in the space $S_A \times S_B$. Assuming that the datasets A and B are identical to those used in the previous, two-pendulum example, the "special" permutation $\tilde{\pi}$ that we have found in the previous subsection will be *identical* to the one we found now. In other words, the presence or absence of the dataset C in the statistical analysis of the Appendix A will not change the permutation $\tilde{\pi}$, although a priori this is mathematically possible. The fact that this does not happen in our thought experiment is a consequence of the nontrivial nature of the dataset (A, B, C).

Since the permutation $\tilde{\pi}$ is independent of the presence of the third dataset, we can use the third dataset to "resolve" the self-intersecting points of the projection of the curve $\mathcal{C}$ to the space $S_A \times S_B$. Namely, the presence of the third observable establishes that we are actually observing a projection of two different points of the curve $\mathcal{C}$ onto the same point in the $S_A \times S_B$ subspace. Specifically, if we have the following two points on the curve $\mathcal{C}$,

$$\begin{aligned} \mathcal{P}_k &\equiv (a_k, b_{\tilde{\pi}(k)}, c_{\tilde{\rho}(k)})\,, \\ \mathcal{P}_m &\equiv (a_m, b_{\tilde{\pi}(m)}, c_{\tilde{\rho}(m)})\,, \end{aligned} \tag{16}$$

such that $a_k \approx a_m$, $b_{\tilde{\pi}(k)} \approx b_{\tilde{\pi}(m)}$ but $c_{\tilde{\rho}(k)} \neq c_{\tilde{\rho}(m)}$, we see that although the points $\mathcal{P}_k$ and $\mathcal{P}_m$ may belong to quite distant parts of the curve $\mathcal{C}$, they will both project to the "same point" in the subspace $S_A \times S_B$,

leading to the apparent intersection. This is again a general feature of mechanical systems — if our curve $\mathcal{C}$ happens to self-intersect anywhere, we can always enlarge the physical system by adding another observable which will "distinguish" between the two appearences of the apparent intersection point, resolving it into two distinct points.

Finally, let us just shortly note that the issue of space-filling curve $\mathcal{C}$ can be eliminated by extending the physical system with an observable x which has a noncompact domain S_X, and can be considered to be monotonically increasing, such as entropy or whatever else is at hand. This means that our big configuration space will be noncompact, and the curve $\mathcal{C}$ will intersect every hypersurface $x = const$ precisely once, sidestepping any issue of space-filling curves.

2.5 The time manifold

The thought experiment described above provides an operational protocol to establish the existence of a 1-dimensional, non-self-intersecting, non-space-filling curve $\mathcal{C}$ as a joint configuration space of a given number of observable mechanical degrees of freedom. The curve $\mathcal{C}$ has all the hallmarks of a manifold (it is a nonempty set with a well-defined topology and dimension), and in order to properly promote it into a manifold, all we need is a set of coordinate charts, i.e., an *atlas*. Assuming for simplicity that $\mathcal{C}$ has a topology of an open line, it is enough to consider a single coordinate chart across the whole curve (as opposed to a circle which requires at least two charts), called *time chart*:

$$T : \mathcal{C} \to \mathbb{R}\,, \qquad (a, b, c, \dots) \mapsto t\,. \tag{17}$$

Being a chart, i.e., a homeomorphism, this map is invertible, giving rise to *parametric equations of motion* for the observables $a, b, c, \dots$,

$$T^{-1} : \mathbb{R} \to \mathcal{C}\,, \qquad t \mapsto (a(t), b(t), c(t), \dots)\,. \tag{18}$$

This map defines $a(t)$, $b(t)$, etc., as functions of the *time coordinate* $t \in \mathbb{R}$, and they satisfy the correlation functions (15). Of course, one can introduce a different chart $T' : \mathcal{C} \to \mathbb{R}$, such that the composition $\phi \equiv T' \circ T^{-1} : \mathbb{R} \to \mathbb{R}$ is a homeomorphism of the real line. One can assume that ϕ is a smooth function, in which case it is a diffeomorphism. It defines the *time reparametrization transformation*, as

$$t' = \phi(t)\,, \tag{19}$$

which keeps the correlation functions (15) invariant. This becomes obvious if we use the composition notation $F_k(a(t), b(t), \dots) \equiv (F_k \circ T^{-1})(t)$, so that

$$F_k \circ T^{-1} = 0 \tag{20}$$

obviously implies

$$F_k \circ T^{-1} \circ \phi^{-1} = 0\,, \tag{21}$$

for all $k = 1, 2, \dots, K$, where ϕ^{-1} is a coordinate transformation from t' to t. Since the curve $\mathcal{C}$ is in fact the joint configuration space of K observables, and it is defined by the equations $F_k = 0$, the invariance of the correlation functions establishes the invariance of the joint configuration space $\mathcal{C}$ under the action of the 1-dimensional group of diffeomorphisms $\mathit{Diff}(\mathbb{R})$. This feature of $\mathcal{C}$ is called the *time reparametrization invariance*, and the group $\mathit{Diff}(\mathbb{R})$ is a subgroup of the larger group of *spacetime diffeomorphisms*, $\mathit{Diff}(\mathbb{R}^4)$, as we shall see in the next section.

Let us finish this section with a remark that the existence of the curve $\mathcal{C}$ in our dataset means that it is impossible to "erase" the information about coincident measurements of the observables. Recall that we were "cutting the photos" to display only individual pendulums, in order to erase the information about the pairing of experiment outcomes. The existence of the curve $\mathcal{C}$ fully recovers that information from our dataset, which means that this information *cannot be erased*. In other words, the information about "conditional measurements" — measuring the position of the pendulum A *under the condition that* pendulum B has some given position — is itself observable, and is encoded in the correlations present in the dataset.

3 Field theory and spacetime

The generalization of the analysis given in the previous section, from a time manifold to a spacetime manifold, is completely straightforward. The only difference is that we need to use fields instead of mechanical systems in our thought experiment.

For the sake of concreteness, let us imagine that we have a fluid, flowing through some big container (a pipe or a river bed). Suppose we have an instrument, called "a probe", which we can insert into the fluid to measure its various properties. The probe has compact spherical shape and is small enough not to perturb the properties and flow of the fluid, while it samples (i.e., performs a coincident measurement of) various observables. When immersed into the fluid and activated, the probe provides the following set of numbers:

- mass-density ρ^m of the fluid,
- pressure p of the fluid,
- temperature T of the fluid,
- charge-density ρ^e of the fluid,
- magnitude $E \equiv \|\vec{E}\|$ of the electric field inside the fluid,
- magnitude $B \equiv \|\vec{B}\|$ of the magnetic field inside the fluid, and
- the angle θ between $\vec{E}$ and $\vec{B}$, specified as $\theta \equiv \vec{E} \cdot \vec{B}/EB$.

In total, the probe measures the 7 observables

$$\rho^m, \quad p, \quad T, \quad \rho^e, \quad E, \quad B, \quad \theta\,. \tag{22}$$

We then perform N measurements, by using N probes. Each probe randomly activates once, performs the measurements, and transmits the measured values wirelessly to our computer. In order to erase any information about "where and when" the measurements took place, the computer does not keep track of anything but the *unordered sets* of N measurements for each observable,

$$\begin{aligned} \rho^m &= \{\rho^m_1, \ldots, \rho^m_N\}\,, \\ p &= \{p_1, \ldots, p_N\}\,, \\ &\vdots \\ \theta &= \{\theta_1, \ldots, \theta_N\}\,. \end{aligned} \tag{23}$$

Each of the 7 observables belongs to its domain, which is either a compact or a noncompact subset of $\mathbb{R}$. Denote them in turn as

$$S_{\rho^m}\,, \qquad S_p\,, \qquad \ldots\,, \qquad S_\theta\,, \tag{24}$$

and introduce the total (kinematic) configuration space as the Cartesian product of all of these,

$$S_{kin} \equiv S_{\rho^m} \times S_p \times \cdots \times S_\theta\,. \tag{25}$$

Next, since we work with 7 observables, introduce 6 arbitrary permutations $\pi_1, \ldots, \pi_6$ of a set of N elements, and construct a $7 \times N$ matrix as

$$\begin{pmatrix} \rho^m_1 & \rho^m_2 & \cdots & \rho^m_N \\ p_{\pi_1(1)} & p_{\pi_1(2)} & \cdots & p_{\pi_1(N)} \\ & & \ddots & \\ \theta_{\pi_6(1)} & \theta_{\pi_6(2)} & \cdots & \theta_{\pi_6(N)} \end{pmatrix}. \tag{26}$$

Each column in this matrix represents one data point in a 7-dimensional scatter plot, which obviously cannot be drawn on paper but is well-defined nevertheless. Since the permutations $\pi_1, \ldots, \pi_6$ can be chosen completely arbitrary, there are in total $(N!)^6$ such matrices, each with its own 7-dimensional scatter plot. Again for obvious reasons, visual inspection of all these plots is not possible, but the statistical analysis from Appendix A should work just fine, provided enough computational power. As a result of the thought experiment, the statistical analysis will provide us with one "special" choice of permutations $\tilde{\pi}_1, \ldots, \tilde{\pi}_6$, for which all data points align themselves nicely along one hypersurface, denoted $\mathcal{M}$, in the big 7-dimensional configuration space S_{kin}. In the limit $N \to \infty$, it satisfies the properties of being a strict measure-zero subset of S_{kin}:

$$\mathcal{M} \subset S_{kin}\,, \qquad \frac{\text{meas}(\mathcal{M})}{\text{meas}(S_{kin})} = 0\,. \tag{27}$$

The hypersurface $\mathcal{M}$ represents the joint configuration space for our physical system, and exhibits the usual properties of *existence*, *self-reinforcement* and *dimensionality*, all established as an experimental result of our thought experiment. The statistically obtained dimension of $\mathcal{M}$, without any a priori reason whatsoever, turns out to be

$$\dim \mathcal{M} = 4\,. \tag{28}$$

Note that this conclusion about the dimensionality, as well as any conclusion based on experimental data, is ultimately contingent on the precision of the equipment used to obtain that data, and may change with the advance of technology and science. See Subsection 4.3 for a detailed discussion of this issue. See also Subsection 4.5 for a detailed discussion of various theoretical proposals to calculate the dimension of spacetime from first principles, such as spinfoam models, string theory, causal set theory, and others.

Also, as before, if the hypersurface $\mathcal{M}$ happens to self-intersect or fills up the whole space S_{kin}, we should simply add additional convenient observables to our dataset, and convince ourselves that the above inconvenient properties dissapear. Also as before, adding additional observables fails to change the overall dimension of $\mathcal{M}$, which persistently keeps being equal to 4, regardless of the number of sampled observables.

If we extend the number of our observables from 7 to K, denoted $a^{(1)}, \ldots, a^{(K)}$, we find a total of $K-4$ correlation functions

$$\begin{aligned} F_1(a^{(1)}, \ldots, a^{(K)}) &= 0\,, \\ &\vdots \\ F_{K-4}(a^{(1)}, \ldots, a^{(K)}) &= 0\,, \end{aligned} \tag{29}$$

whose set of solutions describes a 4-dimensional hypersurface $\mathcal{M}$ in a K-dimensional configuration space.

The only thing left to do at this point is to introduce a set of coordinate charts, i.e., an atlas,

$$\begin{aligned} f: \quad \mathcal{M} &\to \mathbb{R}^4\,, \\ (a^{(1)}, \ldots, a^{(K)}) &\mapsto x \equiv (t, x, y, z)\,, \end{aligned} \tag{30}$$

and its inverse

$$\begin{aligned} f^{-1}: \quad \mathbb{R}^4 &\to \mathcal{M}\,, \\ x &\mapsto (a^{(1)}(x), \ldots, a^{(K)}(x))\,, \end{aligned} \tag{31}$$

which establish the parametric functions $a^{(k)}(x)$ for the observables, which in turn satisfy all correlation functions (29). The hypersurface $\mathcal{M}$, now established as a proper 4-dimensional manifold, is of course called *spacetime*, while the parametric functions $a^{(k)}(x)$ are called *fields living on spacetime*.

One can also introduce a different chart,

$$\begin{aligned} f': \quad \mathcal{M} &\to \mathbb{R}^4\,, \\ (a^{(1)}, \ldots, a^{(K)}) &\mapsto x' \equiv (t', x', y', z')\,, \end{aligned} \tag{32}$$

such that the composition $\phi = f' \circ f^{-1} : \mathbb{R}^4 \to \mathbb{R}^4$ is a homeomorphism in $\mathbb{R}^4$. If ϕ is smooth, it is a diffeomorphism. It defines a *coordinate transformation*

$$x' = \phi(x)\,, \tag{33}$$

which keeps the correlation functions (29) invariant, similar to the 1-dimensional case from Section 2. Since $\mathcal{M}$ is equivalent to the solution of the system (29), and since it is the joint configuration space for the observables, it is invariant under the action of the group of 4-dimensional diffeomorphisms, $Diff(\mathbb{R}^4)$. This feature of $\mathcal{M}$ is called (passive) *diffeomorphism invariance of spacetime*, or *general coordinate invariance*.

Since $\mathcal{M}$ is the joint configuration space for the observables, choosing a particular data point k in spacetime (one of the measured points on $\mathcal{M}$) gives rise to a K-tuple of particular values of the observables. In our example, those are

$$(\rho^m_k, p_{\tilde{\pi}_1(k)}, \ldots, \theta_{\tilde{\pi}_6(k)})\,, \tag{34}$$

and they reconstruct the information about coincident measurements coming from each particular probe — the information we had tried to erase by ignoring any particular order of values in the dataset (23). As in the case of the 1-dimensional time manifold, this information is present in the correlations of the dataset itself, and *cannot be erased*.

As a final point, note that the above operational reconstruction of spacetime is actually one precise implementation of the relationalism paradigm, defining the spacetime manifold using nothing but fields that supposedly live on it. The idea for a mental image "fields do not live on top of spacetime, but on top of each other" (pointed out for example by Rovelli in [9]) can be explicitly realized if we were to take 4 parametric functions $a^{(1)}(x), \ldots, a^{(4)}(x)$, conveniently chosen so that they uniquely specify a spacetime point x in some given chart, i.e., such that one can solve those 4 parametric functions for the coordinates as functions of the observables,

$$x^\mu(a^{(1)}, \ldots, a^{(4)})\,, \qquad \mu = 0, \ldots, 3\,, \tag{35}$$

and then use them to eliminate x from the remaining $K-4$ observables $a^{(5)}, \ldots, a^{(K)}$. In this way one arrives precisely at the $K-4$ correlation functions (29), which actually define the manifold $\mathcal{M}$ using nothing but the information about fields. These correlation functions are the precise technical implementation of the statement that "fields live on fields", in the sense that we can only observe coincidences among fields, as opposed to the the values of fields "at a given spacetime point".

Nevertheless, the fact that we always have precisely $K-4$ correlation functions (that is, for every choice of K) tells us that the set of solutions of those $K-4$ correlation functions, namely the spacetime manifold $\mathcal{M}$, is itself operationally observable, independent of the choice and the amount of the fields one uses to describe it, and despite the diffeomorphism symmetry of those fields. The invariant properties of $\mathcal{M}$ such as dimension and topology are present as correlations in the experimental data, and there is no theoretical account of why they have the values that we observe in the experiment. The 4-dimensionality and simply connected topology are *brute experimental facts*, and are *independent* of the choice, properties and even the very number of the fields we use to measure them (given current science and technology, as mentioned below equation (28) and discussed in detail in Subsection 4.3). Of course, one could attempt to construct a theoretical model which would be able to deduce these correlations among observables from some simpler set of principles. This would implement the concept of the *emergence of spacetime*. So far, however, no such model has ever been successfully constructed.

4 Conclusions and discussion

In this work, we have argued that one can give an operational, model independent experimental protocol whose outcome would be the determination of the dimension and the topology of the time manifold (Section 2) and the spacetime manifold (Section 3). After introducing the thought experiment, we have studied the relevant properties (see Conjecture 2 from Section 2) of the experimental data which give rise to a signal that reflects the observability of the underlying manifold, over which the fields are defined.

In what follows, we will discuss various aspects of the thought experiment that were not discussed in detail in previous sections, but nevertheless deserve to be mentioned and commented on.

4.1 Distinguishing space from spacetime

In Section 2 we have introduced the time manifold by looking at a mechanical system, while in Section 3 we have introduced the spacetime manifold, looking at fields instead. However, one may ask a natural question about the notion of a space manifold, and its differences from spacetime. To that end, let us discuss another illustrative example of the application of our gedanken-eksperiment.

Consider a room with a lamp and some furniture. The experimental apparatus used to perform measurements over this system consists of two cameras separated several centimeters apart. Each camera performs a measurement that provides the following data:

- the polar and azimuthal angles θ and φ of the incident light ray,
- the frequency ν and intensity I of the ray.

Since we have two cameras, in total there are eight observables per measurement. As always, we collect the measured data ignoring any order, and perform the analysis of the protocol, to reach the conclusion that there are five correlation functions between observables, which means that all measurements can be arranged on a 3-dimensional manifold. Intuitively, this camera setup corresponds to the stereoscopic eyesight, that gives one a perception of depth and thus the notion of a 3-dimensional space of the room. One should note that the notion of time is absent from this description, since the scene of the room is static. Hence the resulting 3-dimensional manifold deserves the name *space manifold*.

Nevertheless, one can also consider a room with a lamp and furniture, and additionally a cat moving around inside. If we now perform the same type of measurements of the system with identical cameras as before, we will find only four correlation functions between eight observables, which means that all measurements can be arranged on a 4-dimensional manifold, rather than a 3-dimensional one. It is obvious that the motion of the cat renders the dataset fundamentally more complicated, with less amount of correlation. In other words, the observed 4-dimensional manifold corresponds to spacetime, since in this case the scene of the room is not static anymore.

This example illustrates the dependence of the outcome of the experiment on the properties of the system being observed. It may happen that a system has a high level of symmetry (in the example above, the time-translation invariance), which introduces additional correlations into the dataset and lowers the dimension of the resulting manifold. This is the main way to distinguish space from spacetime — space is in fact spacetime with a certain global symmetry, which renders time unobservable. Similarly, the time manifold from the pendulum example in Section 2 is also spacetime with a global space-translation symmetry (the pendulum swings the same way regardless of its spatial position), which renders space unobservable.

4.2 Quantum-mechanical treatment

It is not completely obvious what would change regarding the outcomes of our thought experiment, if one were to take into account the effects of non-commuting observables, i.e., quantum effects. On one hand, the above analysis makes use of only mutually commuting observables (positions of several pendulums in mechanics, or values of different fields in field theory all mutually commute). This may suggest that our results should not be disturbed by the fact that there exist other observables, which fail to commute with the ones used in the experiment. In principle, one can extend the set of sampled observables up to the so-called *complete set of compatible observables*, without changing anything in the above analysis.

On the other hand, a very precise measurement of a coordinate of a given pendulum may uncontrollably perturb its momentum, so much that the pendulum fails to swing in the usual way, which would introduce a form of an intrinsic noise in the expermental data. In that case, subsequent measurements of the position of the pendulum may fail to be well correlated to the measurements of other observed pendulums, potentially hindering the predictions for the dimension and topology of the proper configuration space. Indeed, whenever the Hamiltonian does not commute with the position operator, a precise position measurement inherently disturbs the system's subsequent dynamics, resulting in potentially unbounded disturbances. This should be taken into account in any generalisation of our gedanken-experiment to the quantum realm.

Additionally, care should be taken to distinguish the single-shot measurement of a single instance of a pendulum (which gives a random result, per QM), from the statistical measurement of an ensemble of identically prepared pendulums (which is probabilistically determined by QM). The uncertainty relations hold for the latter, while in our study we are interested in the former.

The proper quantum mechanical treatment is out of the scope of this paper, and we postpone it for future work.

4.3 Extra dimensions of spacetime

Throughout Section 3, it was claimed that the correlations one ought to find in real experimental data will ultimately support the conclusion $D \equiv \dim \mathcal{M} = 4$. This claim is supported by everyday experience and virtually all experiments ever performed in the history of physics so far. These experiments roughly cover the scales from $10^{-20}\,\mathrm{m}$ (the scale of the current LHC and LIGO experiments), up to $10^{26}\,\mathrm{m}$ (the scale of the observable Universe). There is a further range of scales, from $10^{-20}\,\mathrm{m}$ down to $10^{-35}\,\mathrm{m}$ (the Planck scale), and maybe even beyond that. We currently have no experimental data from this range to either support the result $D = 4$ or disprove it. Eventual access to this data could in principle change the result for D. For example, in the context of string theory [11, 12], one imagines spacetime to have six additional spacelike dimensions, wrapped up into a small Calabi-Yau manifold. If we were to measure various fields at the scale smaller than the size of that Calabi-Yau manifold, our analysis of the data would yield $D = 10$, or 4 plus whatever number of compactified small extra dimensions exist at that scale.

In this sense, the dimension of the spacetime manifold may be a scale-dependent quantity, like running coupling constants in QFT, having different "effective" values at different scales. So far we have no data that would indicate anything other than $D = 4$, but in principle this may change. Related to this,

one may ask if our analysis supports the "running" of D to values smaller than 4 at smaller scales. Values larger than 4 are obviously possible, but smaller values are also possible in principle. Namely, with sufficient resolution, one may notice that what looks like a 4-manifold is (for example) in fact a 1-dimensional curve densely packed up in a space-filling fashion, like some finite iteration of the Peano curve. In this sense, a fundamentally 1-dimensional manifold can look at large scales as a 4-dimensional one. Something along these lines apparently happens in the context of Causal Dynamical Triangulations (CDT) scenario [13, 14], giving rise to $D = 2$ near the Planck scale. That said, note that the CDT results discuss a different concept of a spacetime dimension, namely an effective dimension visible to a random walker. This is in general not equivalent to the notion of topological dimension that we discuss here. Namely, in the context of chaos theory, one typically introduces the notion of a so-called fractal or Hausdorff dimension, which essentially captures the effective dimension visible to a random walker exploring the fractal structure. In general, this notion is different from the ordinary, topological definition of dimension, that we use in this work (see equation (49) from Appendix A). Nevertheless, in cases where the geometric structure under investigation has appropriately 'nice' features, the notions of the Hausdorff and topological dimensions coincide, and give an integer value that corresponds to our intuitive notion of dimension (see [15] for further details). Another context in which one often discusses different effective values of dimension at different scales is Causal Set Theory (CST) [16]. Namely, in this scenario, the fundamental structure describing spacetime is a discrete set of points, which ultimately has topological dimension zero, but in a semiclassical limit gives rise to general relativity and the traditional notion of a 4-dimensional smooth manifold as an effective description on large scales. At the conceptual level, the running of the dimension from 4 to 0 could also be observed using our technique, provided that one can perform measurements at the scale where spacetime discreteness is dominant. Unfortunately, we have no technological capabilities to do this in practice. Nevertheless, the relationship between different dimensions of the manifold at different scales is also well supported theoretically in the CST context, see for example [10].

The result $D = 4$ is also contingent on the choice of observables measured and used to conclude that $D = 4$. In principle, if we were to extend our set of observables, we could find additional *large* dimensions. For example, so far we can detect the presence of dark matter only through its gravitational interaction with regular matter. In addition to that, one could imagine that there are also direct (contact) interactions between dark and regular matter, but that they are obscure enough not to be easily visible (like neutrinos, which interact with other matter only through short-range weak interactions and even weaker gravity). Nevertheless, if we somehow manage to measure these additional observables dependent on dark matter, they may change our statistics even in the IR regime, i.e., at large scales. It may thus turn out that our analysis yields $D = 7$ or $D = 12$ or any number larger than 4. An example of this scenarios are braneworld cosmologies similar to Randall–Sundrum model [17, 18] and similar, where all "standard" observables we can measure happen to be nonzero only on a 4-dimensional submanifold of some target manifold of larger dimension, while the "dark matter" observables would be nonzero even in the bulk of the target manifold, ultimately giving rise to $D > 4$. So far we have not found any such observables, but this may also change in the future.

4.4 Uniqueness objections

Looking at the analysis of the dataset of our thought experiment, one can raise an objection that the analysis of the correlations between observables may fail to give a unique result. Namely, in generic circumstances, any finite dataset constructed by N measurements of K observables can be permuted into $(N!)^{K-1}$ possible arrangements, and for each of them one can evaluate the suitable critical parameter (see Appendix A) that singles out the one "special" permutation. Since the set of permutations is finite, we end up with a finite collections of the values of the critical parameter, and any such collection contains a minimal element. This minimal parameter corresponds to a "special" permutation, since it features the strongest correlation in the dataset, ultimately describing a spacetime manifold. Given this setup, it may turn out that this minimal parameter is not unique, in the sense that several different permutations of the dataset all have this same minimal value of the parameter. In such a case, there are in principle several different possible arrangements of data featuring equally strong correlations, leading to several different possible candidate manifolds.

On one hand, it is not feasible to study this question numerically in practice, since the set of $(N!)^{K-1}$ of all possible permutations is incredibly huge, while permutations featuring the minimal parameter are likely to be a scarce subset of these. This limits us to theoretical arguments that generic datasets feature minimal critical parameter for only one permutation. The main argument provided so far is based on the *self-reinforcement* property, which is unlikely to hold for more than one permutation and its restrictions to arbitrary subsets of data.

On the other hand, it is easy to construct explicit examples which feature more than one "special" permutation, each corresponding to the same minimal value of the critical parameter. This is expected to

happen if a physical system being measured features global symmetries. For example, pendulums are invariant with respect to a global left-right symmetry, in the sense that one can find two "special" permutations, which are mirror images of each other. While these formally give rise to two different time manifolds, these manifolds are represented by two isomorphic datasets, and therefore in fact represent the one and the same time manifold, despite the nonuniqueness. In this sense, nonuniqueness that stems from the existence of global symmetries is benign, and does not invalidate the analysis and conclusions of the thought experiment.

Also, one of the most often discussed examples is the "diagonal" permutation of the dataset, where the values for each observable are sorted in an ascending (or descending) fashion. This leads to a dataset where points are roughly aligned along a 1-dimensional curve connecting two opposite corners of the K-dimensional kinematic configuration space. Being 1-dimensional, this permutation is likely to have a very small value of the critical parameter, possibly the minimal one. Nevertheless, it is not hard to demonstrate that this permutation fails to satisfy the self-reinforcement property, since its restrictions to data subsets will fail to coincide with the full permutation. In other words, as one adds more data to the dataset, the resulting 1-dimensional curve "wiggles", changing its shape for every extension of the dataset. Thus such a permutation has to be excluded from consideration, regardless of the very small value of the critical parameter. We conjecture that the self-reinforcement property will be violated for all permutations featuring small critical parameter, except for one — which can thus be uniquely recongized as "special". However, any potential proof of this conjecture is not available at this time, and requires further study.

4.5 Emergence of spacetime

As we noted at the end of Section 3, in principle one can imagine a theoretical model which does not feature anything like a 4-manifold in its postulated structure. Further, such a model may give us predictions for the values of all possible observables, and we can use it to generate a dataset, study it using our methods, and end up with correlations among observables which give rise to a 4-manifold. If such a scenario happens, one says that spacetime *emerges* from the theoretical model, and that the model predicts the values of spacetime dimension and topology.

Despite many hopeful attempts (usually in the context of quantum gravity), there are no particular theoretical models that have managed to achieve this, even with a wrong prediction for the dimension and topology. The reason for this are two important criteria that such a model must satisfy in order to make a legitimate prediction:

(1) The information about the 4-manifold must not be present in the axiomatic structure of the model. Namely, if we include the 4-manifold structure as an input in the construction of the theory, it should come as no surprise that one can recover that information from the model later on, in various different forms. However, this can be tricky to test, since the information about a 4-manifold can be encoded in a non-obvious, cryptic fashion, and it may be hard to prove that some of the axioms of the model are indeed equivalent to the assumption that there exists a 4-manifold in the theory.

(2) One must demonstrate that the actual dimension of the would-be manifold is explicitly computable from the model. For example, one can try to evaluate a bunch of observables using the model, and then apply our analysis onto that data, in order to obtain $D = 4$. In this sense, the dimension of spacetime would be an explicit consequence of the dynamics of the observables in the model. Alternatively, one may use some other way to calculate the spacetime dimension, but this again needs to be a consequence of dynamics of the observables, giving rise to appropriate correlations in the dataset that would ultimately lead to a 4-manifold. However, it is not enough to handwavingly claim that the model "in principle" predicts these correlations, because they cannot be a generic feature of the model. The correlations must be *explicitly calculated*, or otherwise rigorously proved to exist, and in addition it must be also rigorously proved that there cannot be any further correlations beyond these, since any further correlations would lower the dimension below 4, falsifying the model.

For example, a typical spinfoam model of quantum gravity is constructed as a state sum over the values of the fields living on a 2-complex (the spin foam), providing one with a way to calculate (expectation) values of observables [19, 20]. However, usually by construction, the 2-complex which is used is dual to a triangulation of a 4-manifold. Because of this property, the spinfoam model fails to satisfy the criterion (1) — the information about a 4-manifold is already integrated into the model as an axiom, rather than being a property of the dynamics of the observables. Moreover, if one manages to "fix" this by generalizing the 2-complex structure somehow, so that it fails to be dual to a manifold (for example, by arranging that each dual cell has different topological dimension), there is criterion (2) — one must use the model to explicitly demonstrate that the observables will always feature appropriate correlations (as a consequence of the dynamics of the model) so that they give rise to a 4-manifold, at least in some large-distance limit. No such model has ever been proposed, and no such cal-

culation has ever been performed.

Another example would be bosonic string theory, which is formulated by explicitly assuming the existence of a D-manifold, albeit with an unspecified value for D [11, 12]. Then one uses the dynamics of the model to evaluate 1-loop beta-functions, and from the requirement that these vanish, one obtains a consistency requirement $D = 26$. The explicit assumption of a D-dimensional manifold is an input to the model, and thus already violates criterion (1). Criterion (2) comes very close to being satisfied, but unfortunately it is contingent on the particular choice of ζ-function regularization, making use of the famous "identity"

$$\zeta(-1) \equiv \sum_{n=1}^{\infty} n \equiv 1 + 2 + 3 + \cdots = -\frac{1}{12}\,, \qquad (36)$$

obtained by analytic continuation of the Riemann ζ-function, i.e., postulated as an axiom of the model. This ultimately renders the value of D to be a part of the definition of the theory, completing the argument that criterion (1) is indeed violated.

A third example would be causal set theory [16]. Here, the fundamental structure of spacetime is assumed to be a discrete set of points, without any assumption about 4-dimensional manifold structure. In this sense, such models of quantum gravity indeed seem to satisfy our criterion (1). Moreover, when studying the semiclassical limit of such a model, one can go quite far in obtaining a classical theory of general relativity on a smooth manifold structure, see for example [21]. In this paper, there is even a discussion about the derivation of the spacetime dimension from the fundamental causal set structure, using various dimension estimators and discriminators. However, any explicit calculation involving those estimators that would uniquely provide the result $D = 4$ is still missing. Thus, the so far discussed causal set models of quantum gravity, or otherwise any future proposals, are yet to satisfy our criterion (2).

One can naturally expect to find many other proposals for spacetime emergence *from various theoretical models* throughout the literature. However, in order to take any of these proposals seriously, one must first demonstrate that criteria (1) and (2) are fulfilled, which is highly nontrivial, and likely has not been done for any of the existing proposals.

Let us note that one might also consider another type of emergence, which stems *from raw experimental data*, as opposed to emergence *from a theoretical model*, which was discussed so far in this Subsection. In fact, the approach taken in this work represents an example of this second type of spacetime emergence. In that sense, the spacetime emergence from raw experimental data is precisely a synonym for the notion of the verification of the existence of spacetime, as used in the title. It is important to emphasise that the term "emergence" can thus have quite different meanings, depending on the context, and one should take care in its use.

Acknowledgments

The authors wish to thank Jovan Janjić for help in implementing numerical algorithms in Mathematica, and to Igor Salom, Časlav Brukner, Marcus Huber, Reinhard Werner and Klaus Fredenhagen for fruitful discussions.

This work is funded by FCT/MECI through national funds and, when applicable, co-funded EU funds under UID/5008: Instituto de Telecomunicações. NP acknowledges the FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT000, as well as the FCT project QuantumPrime reference PTDC/EEI-TEL/8017/2020.

MV was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, and by the Science Fund of the Republic of Serbia, grant 7745968, "Quantum Gravity from Higher Gauge Theory 2021" — QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

References

[1] N. Huggett, *Space from Zeno to Einstein: classic readings with a contemporary commentary*, MIT Press, Cambridge (1999).

[2] B. Dainton, *Time and space*, McGill-Queen's University Press, Montreal (2001).

[3] H. G. Alexander, *The Leibniz-Clarke Correspondence*, Manchester University Press, Manchester (1956).

[4] E. Vailati, *Leibniz and Clarke: A Study of Their Correspondence*, Oxford University Press, New York (1997).

[5] L. Sklar, *Philosophy of Physics*, CRC Press, Boca Raton (1992).

[6] R. Carnap, *An Introduction to the Philosophy of Science*, Basic Books Inc., New York (1966).

[7] M. Jammer, *Concepts of Space: The History of Theories of Space in Physics*, Harvard University Press, Cambridge (1954).

[8] J. A. Wheeler, *A Journey into Gravity and Spacetime*, Scientific American Library, New York (1990).

[9] C. Rovelli, *Quantum Gravity*, Cambridge University Press, Cambridge (2004).

[10] J. Butterfield, "En Route to Reduction: Lorentzian Manifolds and Causal Sets", in *Time and Timelessness in Fundamental Physics and Cosmology, eds. S. De Bianchi, M. Forgione and L. Marongiu*, Springer, Cham, Switzerland (2024).

[11] J. Polchinski, *String Theory Vol. I: An Introduction to the Bosonic String*, Cambridge University Press, Cambridge (1998).

[12] J. Polchinski, *String Theory Vol. II: Superstring Theory and Beyond*, Cambridge University Press, Cambridge (1998).

[13] J. Ambjorn, A. Goerlich, J. Jurkiewicz and R. Loll, *Phys. Rep.* **519**, 127 (2012).

[14] J. Ambjorn, A. Goerlich, J. Jurkiewicz and R. Loll, *Int. J. Mod. Phys. D* **22**, 1330019 (2013).

[15] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications (2nd ed.)*, John Wiley and Sons, Hoboken, New Jersey (2003).

[16] S. Surya, *Living Rev. Relativ.* **22**, 5 (2019).

[17] L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 3370 (1999).

[18] L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 4690 (1999).

[19] J. Engle, E. Livine, R. Pereira and C. Rovelli, *Nucl. Phys.* **B799**, 136 (2008).

[20] L. Freidel, K. Krasnov, *Class. Quant. Grav.* **25** 125018 (2008).

[21] F. Dowker and J. Butterfield, "Recovering General Relativity from a Planck scale discrete theory of quantum gravity", `arXiv:2106.01297`.

A Statistical analysis technique

Suppose we are given a dataset of $N \to \infty$ points as K-tuples in some compact volume of a K-dimensional Euclidean space (the noncompact case can be reduced to the compact case by cutting it into countably infinitely many compact pieces, and studying each piece one by one). The volume of the box is known to be V_K. Our task is to verify whether, and to what extent, these N points are "nicely" aligned on some D-dimensional hypersurface in the big K-dimensional space. Obviously, it is assumed that $D < K$, and the notion of "nice" alignment will be defined rigorously below.

A.1 The critical parameter

We proceed as follows. Pick any datapoint x, and construct around it a K-dimensional cube, such that the point x is in the cube's center. Let the length of each edge of the cube be specified as

$$\epsilon = \sqrt[K]{\frac{V_K}{N}}\,, \tag{37}$$

i.e., such that the volume of the cube is

$$V_{cube} \equiv \epsilon^K = \frac{V_K}{N}\,. \tag{38}$$

Suppose first that datapoints are scattered randomly throughout the box. If we construct the same cube around every datapoint, the total volume of all cubes can be represented as:

$$V_{total} = \sum_{n=1}^{N} \frac{(-1)^{n+1}}{n!} \sum_{i_1 \neq \ldots \neq i_n} V_{i_1 \ldots i_n}\,. \tag{39}$$

Here, $V_{i_1 \ldots i_n}$ is the volume of the joint intersection of n cubes labeled by indices $i_1, \ldots, i_n$, and the second sum is evaluated over all nonrepeating values of these indices, which take values from the set $\{1, \ldots, N\}$. Given that all individual cubes are of the same volume, the term for $n = 1$ is

$$V_{n=1} = \sum_{i=1}^{N} V_i = N V_{cube} = V_K\,. \tag{40}$$

Up to an overall minus sign, the remaining terms in (39) are collected into a total *overlap volume*

$$V_{overlap} \equiv -\sum_{n=2}^{N} \frac{(-1)^{n+1}}{n!} \sum_{i_1 \neq \ldots \neq i_n} V_{i_1 \ldots i_n}\,, \tag{41}$$

so that we have

$$V_{total} = V_K - V_{overlap} = \alpha V_K\,. \tag{42}$$

The coefficient $\alpha \in [0,1]$ is another measure of this overlap:

$$\alpha = 1 - \frac{V_{overlap}}{V_K} \,. \tag{43}$$

In the main text, we call α the *critical parameter*.

Specifically, if the datapoints are very uniformly distributed throughout the box, their corresponding cubes will have very little overlap volume, $V_{overlap} \approx 0$, so that the volume covered by all cubes will be approximately equal to the volume of the box, V_K, and $\alpha \approx 1$. On the other hand, one can suppose that the datapoints are all aligned along some D-dimensional hypersurface $\mathcal{M}$, whose total D-dimensional volume is V_D, while its K-dimensional volume is of course of measure zero. Now, if all N datapoints are aligned on $\mathcal{M}$, their cubes will overlap quite a lot, giving $\alpha \approx 0$ in the limit $N \to \infty$.

One can estimate how close α is to zero in the following way. Construct a cube around every point lying on $\mathcal{M}$. The intersection between $\mathcal{M}$ and the cube will be again a D-dimensional hypersurface, while the cubes will "protrude out" of $\mathcal{M}$ into $K-D$ orthogonal directions. We end up with a "thick hypersurface", having thickness ϵ^{K-D}. Since by assumption all N points are lying on $\mathcal{M}$, one can approximate the value of V_{total} of this thick hypersufrace as

$$V_{total} \approx \epsilon^{K-D} V_D = \left(\frac{V_K}{N}\right)^{\frac{K-D}{K}} V_D \,. \tag{44}$$

This gives us the following estimate for α:

$$\alpha(N) \equiv \frac{V_{total}}{V_K} \approx \frac{V_D}{(\sqrt[K]{V_K})^D} N^{\frac{D}{K}-1} = const \cdot N^{\frac{D}{K}-1} \,. \tag{45}$$

Since $D < K$, we see that in the limit $N \to \infty$ we have $\alpha \to 0$. Pictorially, the "thickness" ϵ^{K-D} of the hypersurface gradually shrinks as N grows, due to (37), so that in the limit $N \to \infty$ we have $\epsilon \to 0$, and our thick hypersurface becomes asymptotically infinitely thin. This is all illustrated for the case $K = 2$ in the diagrams given in Figure 6. On the other hand, if we take a random permutation of the second coordinates,

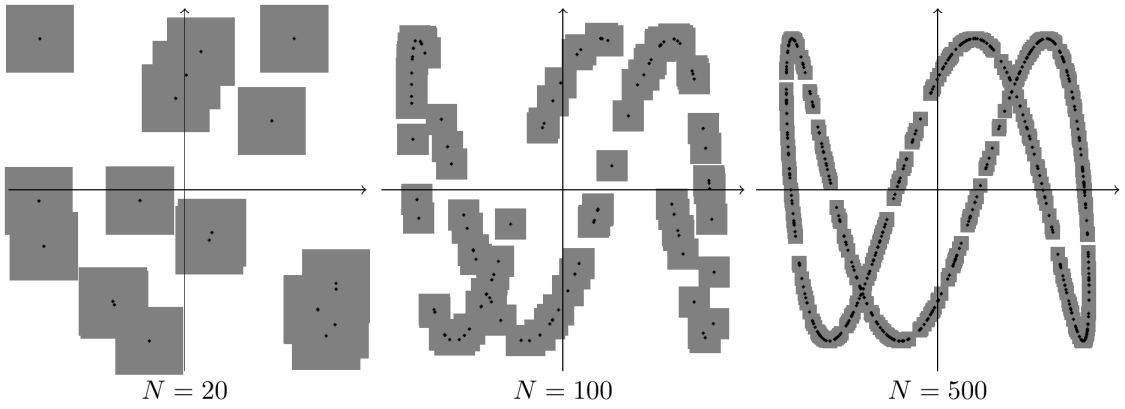


Figure 6: Successive approximations of the datasets, consisting of N points, with a manifold-like structure.

the resulting diagrams look as the ones from Figure 7. We can observe that, as N grows, the area covered by the gray squares becomes smaller in the first case, while it remains approximately constant in the second (permuted) case. In the limit $N \to \infty$ the gray area drops to zero in the first case, while it remains constant in the second case.

A.2 The asymptotic behavior

Let us now formalize the procedure for the statistical analysis. Given the dataset of N points and the volume V_K of the big space, a computer can actually evaluate V_{total} by explicit numerical evaluation of the overlap $V^{(n)}_{overlap}$ between the cube constructed around each point n and all surrounding cubes. This means that $\alpha(N)$ can be explicitly evaluated, for each N. Provided that we have evaluated it for a number of different values of N, we can discuss its asymptotic structure in the limit $N \gg 1$ (which is formally equivalent to $N \to \infty$), as follows:

$$\alpha(N) = c_K + \sum_{n=0}^{K-1} c_n \frac{1}{(\sqrt[K]{N})^{K-n}} + \mathcal{O}\left(\frac{1}{(\sqrt[K]{N})^{K+1}}\right) + \mathcal{R}(N) \,. \tag{46}$$

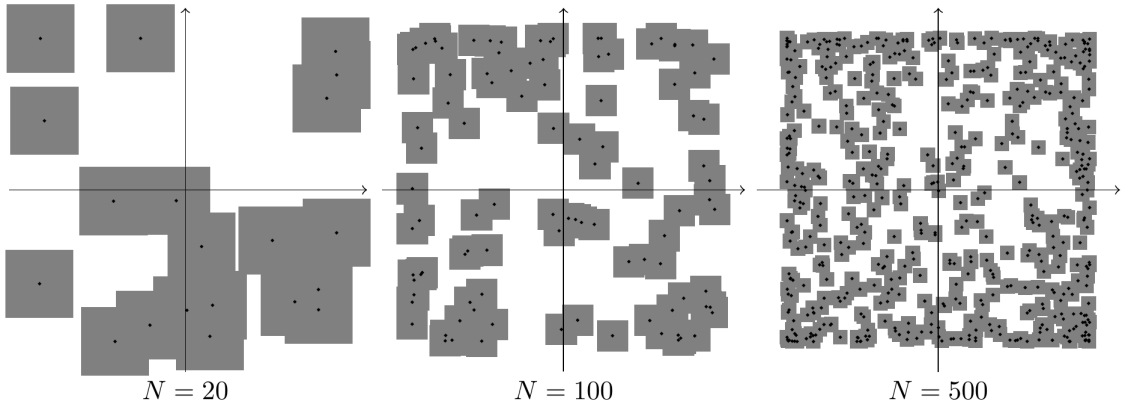


Figure 7: Successive approximations of the datasets, consisting of N points, with a non-manifold-like structure.

Here we have expanded $\alpha(N)$ into an asymptotic power series, from the order $1/\sqrt[K]{N}$ up to the order $\mathcal{O}\left(1/(\sqrt[K]{N})^{K+1}\right)$, while the remainder-term $\mathcal{R}(N)$ captures any behavior of $\alpha(N)$ which does not have the form of this power series — terms like $\log N$, or N^N, or various other types of asymptotics.

Now we define that the *datapoints are "nicely" aligned along a D-dimensional hypersurface* if and only if in the limit $N \gg 1$ the asymptotics has the following form:

$$\mathcal{R}(N) = 0\,, \qquad c_K = 0\,, \qquad c_n = \begin{cases} c_D & n = D\,, \\ 0 & n \neq D\,, \end{cases} \qquad \forall n = 0, \ldots, K-1\,. \tag{47}$$

This is actually a statement that $\alpha(N)$ has the asymptotics of the form (45), while the single nonzero coefficient c_D is proportional to the D-dimensional volume of the hypersurface,

$$c_D \equiv \frac{V_D}{(\sqrt[K]{V_K})^D}\,. \tag{48}$$

Moreover, in this case one can solve (45) for D, which gives us the *statistically defined dimension of the hypersurface*,

$$D = K\left[1 + \lim_{N\to\infty} \frac{\log \alpha(N)}{\log N}\right]\,. \tag{49}$$

Thus, if the asymptotics of $\alpha(N)$ is "nice" in the above sense, the quantity D will be a positive integer smaller than K (since $\log \alpha(N)$ is negative), and it can be explicitly computed, asymptotically for ever larger N.

Finally, if we have a bunch of available datasets (for example, $(N!)^6$, as we do in the main text), we can calculate the asymptotic form (46) for each. Among all these, the one dataset that converges most efficiently toward (45) is called "special", and the corresponding value of D calculated from (49) is called its *dimension*. It defines a hypersurface $\mathcal{M}$ as a strict subspace of the box of volume V_K, and is moreover of measure zero compared to the box, since

$$\frac{\text{meas}(\mathcal{M})}{\text{meas}(\text{box})} \equiv \lim_{N\to\infty} \frac{V_{total}}{V_K} = \lim_{N\to\infty} \alpha(N) = 0\,, \tag{50}$$

according to the asymptotics (45).

At the end, we note that we have implicitly also used the *self-reinforcement* property of the dataset, in the sense that the limit $N \to \infty$ actually exists, i.e., that instead of the dataset of N points we can use the dataset of $N + M$ points, such that the asymptotic properties (47) are maintained as we pass from one dataset to another. This is necessary for the limit $N \to \infty$ to actually make sense, in this context (see also [10] for further discussion of the discrete-to-continuum limit).

In light of the asymptotic formulas (46) and (47), the three properties from the main text, namely (a) *existence*, (b) *self-reinforcement* and (c) *dimensionality* mean in turn that (a) $c_K = \mathcal{R}(N) = 0$, (b) the limit $N \to \infty$ is well-defined, and (c) only one of the coefficients c_n is different from zero. Finally, the fourth property, namely (d) *topology*, can be determined from the overlapping K-dimensional cubes constructed around the datapoints. Specifically, in the limit $N \to \infty$ the total overlap volume $V_{overlap}$ must remain finite, precisely because V_{total} falls to zero. This means that, despite the fact that each K-dimensional cube shrinks to zero, its overlap with neighboring cubes will remain finite. This property enables us to define *a basis of open sets* on our would-be manifold $\mathcal{M}$ as the D-dimensional projections of all K-dimensional cubes (of volume V_{cube}), giving rise to a *topology on* $\mathcal{M}$.

A.3 Alternative technique for evaluating the critical parameter

The algorithm for calculating $\alpha(N)$, while conceptually clear, turns out to be a bit inefficient for large datasets, due to the complexity of calculating the overlap volumes of neighboring cubes. The algorithm starts to choke already for $K = 2$ and $N = 100$ on a typical desktop computer. To remedy this, instead of using (43) to calculate the parameter α, one should instead split the whole box into a convenient grid of cells, and then count the cells which contain at least one datapoint. The total volume of the nonempty cells should then be similar to the total volume of the overlapping cubes from the original approach.

To implement this idea, we first choose the number of cells along each axis of the box to be

$$\left\lfloor \sqrt[K]{N} \right\rfloor , \tag{51}$$

so that the total number of cells is

$$\left\lfloor \sqrt[K]{N} \right\rfloor^K \leq N . \tag{52}$$

This is the biggest integer smaller than N such that the K-dimensional box can be divided into that many cells. The size of each cell in the n-th direction x_n is then given as

$$\tilde{\epsilon}_n = \frac{x_n^{\max} - x_n^{\min}}{\left\lfloor \sqrt[K]{N} \right\rfloor} , \tag{53}$$

so that the total volume of each cell is

$$\tilde{V}_{\text{cell}} = \prod_{n=1}^{K} \tilde{\epsilon}_n = \frac{V_K}{\left\lfloor \sqrt[K]{N} \right\rfloor^K} \geq \frac{V_K}{N} \equiv V_{\text{cube}} , \tag{54}$$

see (38). Given this arrangement, we can define a new parameter $\tilde{\alpha}$, in analogy to (42), as

$$\tilde{V}_{\text{total}} \equiv p\tilde{V}_{\text{cell}} = \tilde{\alpha} V_K , \tag{55}$$

where p is the number of cells which contain at least one datapoint. We thus arrive at a new parameter,

$$\tilde{\alpha} = p\frac{\tilde{V}_{\text{cell}}}{V_K} = \frac{p}{\left\lfloor \sqrt[K]{N} \right\rfloor^K} . \tag{56}$$

Note that we have used the "floor" function in (51) in order to avoid underestimating the volume (55), and thus avoid underestimating $\tilde{\alpha}$ and consequently D.

Compared to (43), parameter $\tilde{\alpha}$ can be calculated much more efficiently, since it boils down to counting the number of nonempty cells p, which is way faster than evaluating all the overlapping volumes of cubes. Indeed, for the case $K = 2$ and $N = 10^4$, the algorithm takes around 10 s to evaluate $\tilde{\alpha}$ on the same hardware as before.

Intuitively, in the limit $N \to \infty$ one expects that $\tilde{\alpha} \approx \alpha$, at least in the case of "nice" alignment of datapoints (as defined by (47)). This establishes that in practical simulations we can calculate the statistical dimension D of the dataset (see (49)) using $\tilde{\alpha}$ instead of α, which is numerically much more efficient. Indeed, this is also confirmed by explicit numerical calculations on several different example datasets.

On the other hand, using the grid-like construction above obscures the notion of basis of open sets, since cells in the grid never overlap, and therefore one cannot use this construction to introduce the topology of the manifold. In this sense, while the grid-like construction of $\tilde{\alpha}$ is numerically more efficient, the overlapping-cubes construction of α gives us the information about topology and is thus conceptually more useful.

arXiv:2305.14252v3 [quant-ph] 15 Jan 2024

Quantum Kolmogorov complexity and quantum correlations in deterministic-control quantum Turing machines

Mariano Lemus[1,2], Ricardo Faleiro[1,2], Paulo Mateus[1,2], Nikola Paunković[1,2], and André Souto[2,3,4]

[1]Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av.Rovisco Pais, 1049-001 Lisboa, Portugal
[2]Instituto de Telecomunicações, Av.Rovisco Pais, 1049-001 Lisboa, Portugal
[3]Lasige, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
[4]Departamento de Informática,Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

This work presents a study of Kolmogorov complexity for general quantum states from the perspective of deterministic-control quantum Turing Machines (dcq-TM). We extend the dcq-TM model to incorporate mixed state inputs and outputs, and define dcq-computable states as those that can be approximated by a dcq-TM. Moreover, we introduce (conditional) Kolmogorov complexity of quantum states and use it to study three particular aspects of the algorithmic information contained in a quantum state: a comparison of the information in a quantum state with that of its classical representation as an array of real numbers, an exploration of the limits of quantum state copying in the context of algorithmic complexity, and study of the complexity of correlations in quantum systems, resulting in a correlation-aware definition for algorithmic mutual information that satisfies symmetry of information property.

1 Introduction

Kolmogorov complexity is a mathematical formulation to capture the intuitive idea of the amount of information an individual object (usually a string) has. This notion was independently proposed by Kolmogorov in 1965 [12], Solomonoff in 1964 [30] and Chaitin in 1966 [7]. It defines the amount of information of an object as the number of minimal instructions to be given to a machine to reproduce that object. If the object is a string x, and the machine is a Universal Turing machine U, the Kolmogorov complexity of x is the length of the shortest program that running in U as input, outputs x. We refer to the book by Li and Vitányi [14] for a comprehensive reading of this theme.

Several aspects must be considered to extend the notion of complexity to the quantum realm, as, in this context, the objects are quantum states, and the set of potential operations is not discrete. In [35], Vitányi identifies three approaches for the quantitative definition of information in a quantum state. The first one is the description (using classical bits) of the physical apparatus that outputs the given state. In this context, the continuous set of quantum states can be approximated by Cauchy sequences of states that can be directly outputted by the machine. The second approach consist on the *qubit* description of the quantum machine outputting the state, which upper bounds the information in a given quantum state by the logarithm of the dimension of the associated Hilbert space. The third approach considers

Mariano Lemus: mlemus@tecnico.ulisboa.pt, Corresponding author
Ricardo Faleiro: ricardofaleiro@tecnico.ulisboa.pt
Paulo Mateus: paulo.mateus@tecnico.ulisboa.pt
Nikola Paunković: npaunkov@mathtecnico.ulisboa.pt
André Souto: ansouto@ciencias.ulisboa.pt

the information content in a quantum state to be the classical information content of the list of real numbers involved in a (fixed) mathematical expression for the state. Furthermore, there are additional aspects to consider, including how to handle the distinction between a machine outputting a state or a high fidelity approximation of it; and the issue of directly addressing mixed states, and therefore define the machine through density operators instead of pure states.

In the last few decades, several definitions and approaches have been proposed in the literature that explore all the aforementioned possibilities. Svozil [32], in his pioneer work, defined the quantum Kolmogorov complexity of a state as the length of the minimum classical description in a quantum Turing machine (in the machine model defined in [8]) of that state. In [34], Vitányi proposed the definition of the (quantum) Kolmogorov complexity of a pure quantum state $|\varphi\rangle$ as the minimum value of the sum of two parts: (i) the length of a classical program computing $|\psi\rangle$, an approximation of $|\varphi\rangle$, and (ii) the negative log-fidelity of the approximation of $|\psi\rangle$ to $|\varphi\rangle$.

These two first approaches considered classical descriptions for quantum states. In [6], the authors considered the possibility of using a quantum program (encoded in qubits instead of bits) to define the complexity similarly to Vitányi's. In their work, Berthiaume, van Dam, and Laplante defined the Kolmogorov complexity of a qubit string as the length of the shortest quantum bit string that, given to a quantum universal Turing machine, produces the qubit string with high fidelity. Its properties were further studied in [23, 24].

Based on a generalization of the notion of universal semi-measure, Gács [9] proposed to define Kolmogorov complexity based on density matrices. Later, in [21], the Kolmogorov complexity of a pure quantum state $|\varphi\rangle$ was defined as the length of the shortest description of a quantum circuit capable of producing such state (or an approximation of it). More recently, a new approach to define Kolmogorov complexity based on a particular type of quantum Turing machines, called deterministic controlled Turing machines was proposed [16]. In this approach, a description of a quantum state $|\varphi\rangle$ is given by a sequence of quantum gates, which transforms an initial reference state of the quantum tape into the state $|\varphi\rangle$.

Although (classical) Kolmogorov complexity has been successfully applied to several fields ranging from security, computational complexity, bio-medicine, etc [1, 2, 31], corresponding applications of its quantum versions are yet just a few. For example, in [21] the authors establish the ground to characterize quantum entanglement via Kolmogorov complexity. In [22] Mora and Briegel and Kraus studied the application of quantum Kolmogorov complexity to thermodynamics and complexity theory. In [3] and [17] the authors explore the connection between (quantum) Kolmogorov complexity and entropy applied to Brudno's theorem and information-disturbance theorem, respectively. Another example of application is in cryptography, in [18], the authors use (classical) Kolmogorov complexity to derive the security of BB84 quantum key distribution [4]. Recently, in [28], algorithmic information theory has been used in genomics applications.

The interconnection between Kolmogorov Complexity Theory and Information Theory is well-known and extensively studied in the literature [10, 33]. Concepts in one theory can find analogous versions in the other. One of the most important concepts in Information Theory is the notion of mutual information, which captures the inherent dependency of how much information is needed to explain one object having another object as a starting point. One of the most significant characteristics of (classical) mutual information is that it is symmetric, that is, $I(X, Y) = I(Y, X)$. On the other hand, for algorithmic mutual information such equality holds up to a logarithmic term of the size of the objects [13, 33, 36]. The concept is briefly mentioned in the quantum Kolmogorov case in [34]. In the context of quantum Kolmogorov complexity theory, the study of quantum algorithmic mutual information can help understand and quantify the complexity of the correlations between parts of multipartite quantum systems, since these systems can be inherently correlated. It is well-known that the correlations in quantum systems can be stronger than those present in classical systems due to the presence of entanglement [11] or the more general quantum discord [20]. Understanding quantum correlations and their relation to algorithmic mutual information measured in terms of a quantum version of Kolmogorov complexity is important, as it opens the possibility of quantifying the amount of manipulation required to create such correlations between two systems. Consequently, this connection can be useful in analyzing complexity in quantum computing, leading to more efficient algorithms, or in relating security and complexity in quantum cryptography protocols, for example.

1.1 Overview of results

The main results of this work can be summarized as follows:

- We present a natural generalization of the *deterministic-control quantum Turing machine* (dcq-TM) model, originally introduced in [16], to allow for mixed state inputs and outputs.
- We define the set of dcq-computable states as the set of states which can be approximated with arbitrary precision using a dcq-TM.
- We define the (conditional) Kolmogorov complexity K of quantum states in the dcq-TM model and show that it is machine independent up to a constant.
- Using the concept of *classical representation* of a quantum state to refer the array of real numbers describing its density matrix in the computational basis, we show that the defined set of dcq-computable states coincides with the set of states with computable classical representations. Moreover, we note that the algorithmic complexity of a quantum state and that of its classical representation are equal up to a constant.
- To contrast with the above result, we compare the properties of quantum states and their classical representations when used as a resource. In this context, we show that the representation contains more descriptive information when used as a resource for computation. We then further explore this difference by reducing the problem of computing an n-qubit state ρ to the problem of computing a *duplicate* of ρ given access to one copy of it, resulting in a relation between the information in a single copy of a quantum state versus the information in two copies of itself. We conclude that for the majority of quantum states, cloning them is essentially as hard as to create them. The result can be interpreted as an algorithmic information version of the no-cloning theorem.
- Finally, we explore the concept of algorithmic mutual information for dcq-computable states. We show that the analogous version of the classical chain rule is not satisfied for general bipartite states and interpret this discrepancy to be caused by the presence of correlations among the two subsystems. We propose an alternative generalization of the chain rule for quantum systems and show that it is indeed satisfied under our definition of K. Using this result, we present a correlation-accounting definition for algorithmic mutual information and show that it satisfies the symmetry of information property.

This paper is organized as follows: In Section 2, we briefly go through some important definitions on classical Kolmogorov complexity and computability of real numbers. In Section 3, we describe the dcq-TM machine model and present definitions for classical representations of quantum states, as well as classical simulators for quantum machines. In Section 4, we define the concept of computable quantum states and compare it to the set of states with computable classical representations. In Section 5, we define the concept of Kolmogorov complexity in the dcq-TM model and study some comparisons with the complexity of the associated classical representations, culminating in a dcq-TM version of no-cloning theorem. Later, in Section 6 we study potential definitions of algorithmic mutual information with the intent of quantifying quantum correlations, resulting in a generalization of the chain rule and an alternative definition for mutual information. Finally, in Section 7, we summarize and discuss the results, as well as point to further research directions on the topic.

2 Preliminaries

2.1 Kolmogorov complexity

In this subsection we briefly review some basic properties of the standard classical algorithmic complexity K in order to better motivate the definitions in the following sections. Classically, the Kolmogorov complexity of a string x is defined as the length of the shortest program that produces x when given to a classical universal Turing machine.[1] We refer the reader to the book of Li and Vitányi [14] for a complete study on this topic.

[1] In this work we focus on inputs to the Turing Machines which need not be prefix-free. In the literature, this is sometimes called "plain" Kolmogorov complexity.

Definition 2.1. *(Kolmogorov complexity)*
Let $x, y \in \{0,1\}^*$ be two strings and U a classical universal Turing machine. The Kolmogorov complexity of x given y is defined as:

$$\mathrm{K}(x \mid y) = \inf_p \{|p| : \mathsf{U}(p, y) = x\}. \tag{1}$$

The default value for y, the auxiliary input for the program p, is the empty string. In order to avoid overloaded notation, in those cases we typically drop this argument in the notation. Notice that Kolmogorov complexity is machine independent in the sense that we can fix a universal Turing machine U whose program size is at most a constant additive term worst than any other machine, and the running time is, at most, a logarithmic multiplicative factor slower than in any other machine (see Theorem 7.1.1 from [14]).

In information theory, one of the fundamental quantities is mutual information. There is an analogous version for Kolmogorov complexity that we will call *algorithmic mutual information*, and is given by

$$\mathrm{I}_{\mathrm{K}}^{(1)}(x, y) = \mathrm{K}(x) - \mathrm{K}(x \mid y). \tag{2}$$

In contrast with standard (Shannon) mutual information, this quantity is independent of any probability distribution and is concerned only with how much of the information in the string y can be used to describe x. In information theory, one useful result is the chain rule, which allows us to express the entropy of a joint distribution in terms of its conditional and marginal distributions. For Kolmogorov complexity we have,

Property 2.1. *(Chain rule 1)*
For all strings x and y in $\{0,1\}^*$,

$$\mathrm{K}(x, y) = \mathrm{K}(x) + \mathrm{K}(y \mid x) + O(\log(\mathrm{K}(x, y))). \tag{3}$$

Using Equation (3) in combination with Equation (2) leads to a result called *symmetry of information*, which states that

$$\mathrm{I}_{\mathrm{K}}^{(1)}(x, y) = \mathrm{I}_{\mathrm{K}}^{(1)}(y, x) + O(\log(\mathrm{K}(x, y))). \tag{4}$$

The chain rule also allows us to relate $I_{\mathrm{K}}^{(1)}$ with another quantity, which can also be understood as a measure of mutual descriptive information

$$\mathrm{I}_{\mathrm{K}}^{(2)}(x, y) = \mathrm{K}(x) + \mathrm{K}(y) - \mathrm{K}(x, y). \tag{5}$$

Analogous expressions for $\mathrm{I}_{\mathrm{K}}^{(1)}$ and $\mathrm{I}_{\mathrm{K}}^{(2)}$ exist in classical information theory for random variables X, Y by replacing the Kolmogorov complexity K by the Shannon entropy function H. In classical information theory those quantities are proven to be equivalent [29]. In the case of algorithmic information theory, they are related by

$$\mathrm{I}_{\mathrm{K}}^{(1)}(x, y) = \mathrm{I}_{\mathrm{K}}^{(2)}(x, y) + O(\log(\mathrm{K}(x, y))). \tag{6}$$

Let $n = |x| + |y|$, by noting that $\mathrm{K}(x, y) \leq \mathrm{K}(xy) + O(\log(n))$ we can slightly modify the chain rule to a form that will be later more suitable for generalizing to the quantum case [14]:

Property 2.2. *(Chain rule 2)*
For all strings x and y in $\{0,1\}^*$ with $n = |x| + |y|$,

$$\mathrm{K}(xy) = \mathrm{K}(x) + \mathrm{K}(y \mid x) + O(\log(n\,\mathrm{K}(xy))) \tag{7}$$

$$= \mathrm{K}(x) + \mathrm{K}(y \mid x) + O(\log(n)). \tag{8}$$

Even though Eq. (8) is the simplest form of the classical chain rule, we want to highlight Eq. (7) because it will allow us to better compare it with its quantum counterpart later on.

2.2 Computability of real numbers

In this section, we briefly review the notions associated with the computability and algorithmic complexity of real numbers. Throughout the paper, we shall need to encode tuples of strings into a single string. Therefore, we denote by $[x_1, \ldots, x_m]$ the encoding of the m-tuple $(x_1, \ldots, x_m)$ of binary strings into a single binary string. Fix a one-to-one encoding of natural numbers in strings and encoding of rational numbers by assigning the information of sign, numerator, and denominator to the different components of a tuple.

Definition 2.2. *A real number r is said to be computable if there exists a (classical) Turing machine T which on any input $k \in \mathbb{N}$ outputs $q_k \in \mathbb{Q}$, such that*

$$|r - q_k| \leq \frac{1}{k}. \tag{9}$$

In this case, we say that T computes r.

Computable real numbers can be encoded effectively in natural numbers. Given an enumeration of the set of Turing machines, we will consider that a computable real number is encoded by any index that represents a Turing machine that will output the approximation of it according to Def. 2.2. A tuple of real numbers $(r_1, \ldots, r_n) \in \mathbb{R}^n$ is computable if all its components r_i are computable. This encompasses complex numbers (ordered pairs of real numbers) and, notably, density matrices (arrays of complex numbers). We can proceed now to define the Kolmogorov complexity of a real number by identifying it with the length of a program that produces a sequence that converges to it according to Equation (9):

Definition 2.3. *Let r be a real number and U a universal classical Turing machine. The Kolmogorov complexity of r is given by*

$$\mathrm{K}(r) = \inf_p \{|p| : \mathsf{U}_p \text{ computes } r)\}, \tag{10}$$

where $\mathsf{U}_p(x) = \mathsf{U}(p, x)$.

The notion Kolmogorov complexity generalizes straightforwardly to arrays of real numbers and matrices as expected.

Remark. *Since the set of computable numbers contains $\mathbb{N}$, some ambiguity may arise when talking about the complexity of a natural number, as both Def. 2.1 and Def. 2.3 may apply (depending on the chosen encoding), but they in general do not coincide. In this document, unless stated otherwise, whenever a number x is understood to take values only in $\mathbb{N}$ (index in an enumeration, counter in a* for *loop, etc.) we assume that $\mathrm{K}(x)$ is given by Def. 2.1. Otherwise, we assume its complexity is given by Def. 2.3.*

3 Definition of the machine

3.1 Deterministic control quantum Turing machine (dcq-TM)

A dcq-TM is a completely deterministic machine in the sense that the flow of computation is entirely classically controlled, but it is also capable of having quantum states as inputs and outputs. The following description of a dcq-TM is based on the one defined in [16] with the added feature that the machine is also able to handle general mixed states, and not just pure states. The machine is understood as having at least two infinite tapes, one classical and one quantum, with their respective classical and a quantum heads[2]. The distinctive feature of the machine is that it has a deterministic control over the computation happening on the quantum tape – this control is only dependent on the set of internal states (Q) of the dcq-TM, which are classical, and the contents of the classical tape. Thus, the set of internal states and the classical input specify the dcq-TM's computational dynamics entirely, and the quantum tape is a

[2]Please note that our "quantum head" is in fact a classical head that operates over the quantum tape. In this sense, it evolves according to a completely classical-deterministic transition function, as opposed to the Deutsch and Bernstein-Vazirani models [5, 8] where the quantum head evolves unitarilly and can be found in superpositions of its classical internal states.

"work only" tape whose contents do not influence the computation. The quantum tape is effectively an infinite quantum memory, composed by quantum cells, each associated with a two dimensional (qubit) Hilbert space $\mathbb{C}^2$. The classical control of the dcq-TM commands the quantum head in a deterministic way. It chooses where the quantum head moves to and picks, from a pre-chosen universal set $\mathbb{U}$ (which without loss of generality is assumed to be finite), which unitary gates to apply and to which cells of the quantum tape to apply them to. It is important to stress that the unitary transformations in $\mathbb{U}$ must be described only with computable complex numbers, otherwise the machine would compute super-Turing distribution/functions. The machine is rigorously defined as follows,

Definition 3.1. *A deterministic-control Quantum Turing machine or dcq-TM is defined by the pair, (Q, δ), where Q and δ are, respectively, the finite set of control states containing at least the two distinct states q_s (starting state) and q_h (halting state) and the transition function of the computation. The transition function is,*

$$\delta : Q \times \mathbb{A} \to \mathbb{U} \times \mathbb{D} \times \mathbb{A} \times \mathbb{D} \times Q, \tag{11}$$

where $\mathbb{A} = \{0, 1, \square\}$ is the alphabet of the classical tape, $\mathbb{U}$ is a universal set of unitary operators that can be applied to the quantum tape, and $\mathbb{D} = \{\mathsf{L}, \mathsf{N}, \mathsf{R}\}$ is set of possible head displacements (left, none, right).

3.1.1 Anchor cell and input/output schemes

In order to have the quantum and classical inputs/outputs properly defined in the dcq-TM we introduce the concept of an *anchor cell*. The anchor cell serves the purpose of specifying a global reference point for each tape, in order to identify unambiguously the inputs and outputs of the machine. For a given tape, we specify the anchor cell (on that tape) to be the first cell before the input starts i.e if we index the input cells with positive numbers $i > 0$, the anchor cell will always be the cell with index $i = 0$. The input and the output schemes of the dcq-TM presented here are slightly different from the ones in the definition of [16], where there is no *anchor cell* defined, and the machine is not aware of the size of its quantum input/output. In the spirit of generalizing the model for general mixed states and more naturally defining inputs and outputs that may start or end with $|0\rangle$ states. We introduce this alternative definition from which it is straightforward to recover all the results already established regarding the dcq-TM.

Inputs – We say the machine received as input the pair $(x; \rho_{\langle 1,n\rangle})$ if its internal state is q_s, its classical and quantum heads are parked in their respective anchor cells, and the tapes' contents are as shown in Figure 1. The classical input x is the string written between the anchor and the first blank after the anchor. At the end of the classical input, we insert a blank cell to distinguish strings x and n, where the latter specifies the quantum input in the quantum tape. The quantum input $\rho_{\langle 1,n\rangle}$ is the physical quantum state present in the set of cells between the anchor and the cell specified by string n.

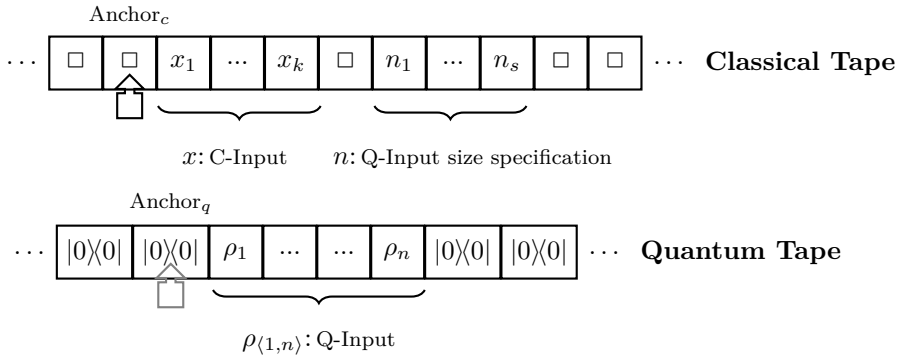


Figure 1: **Starting configuration of the dcq-Turing machine with a classical and quantum input.** $x = x_1 \cdot x_2 \cdots x_k$ is the classical input, while $\rho_{\langle 1,n\rangle}$ is the quantum input "written" on n cells of the quantum tape, where ρ_i is the local state of the $i-$th cell, represented by $\boldsymbol{\rho}_i = \mathrm{Tr}_{(1,\ldots,i-1,i+1,\ldots,n)}[\boldsymbol{\rho}_{\langle 1,n\rangle}]$. The string $x\square n$ is surrounded by blank symbols extending to infinity in both directions. The quantum input is surrounded by cells in the local zero state $|0\rangle\langle 0|$ extending to infinity in both directions.

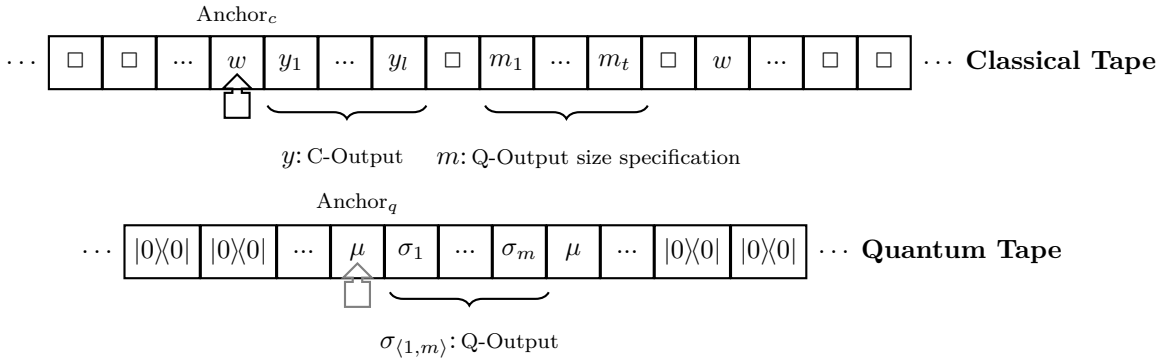


Figure 2: **Final state of the dcq-Turing machine with a classical and quantum outputs.** $y = y_1 \cdot y_2 \cdots y_r$ is the classical output, while $\sigma_{\langle 1,s\rangle}$ is the quantum output, where σ_i is the local state of the $i-$th cell, represented by $\boldsymbol{\sigma}_i = \mathrm{Tr}_{(1,\ldots,i-1,i+1,\ldots,n)}[\boldsymbol{\sigma}_{\langle 1,s\rangle}]$. The size (in number of qubits) of $\sigma_{\langle 1,s\rangle}$ is specified by the string s in the classical tape. The working cells, which are used during the computation but are not part of the output in the classical and quantum tape are denoted by w and μ, respectively.

Outputs – We say the machine produced as output, $(y;\sigma_{\langle 1,s\rangle})$, if the machine is in the halting state $\mathsf{q_h}$ and the classical and quantum tapes are in the configuration shown in Figure 2. The classical output y is defined as the string written in the set of cells that are between the anchor and first blank to the right of the anchor. The quantum output specification m is defined as the string between the first and the second blanks appearing after the anchor. Finally, the quantum output is defined as the physical quantum state present in the set of cells between the anchor and the cell specified by string m.

Although there are many ways to evaluate functions that have multiple classical inputs/outputs using the above configuration, for the purposes of the examples of this paper it is convenient to have a unified encoding to use whenever we refer to algorithms that use a single string x to encode a tuple $(x_1,\ldots,x_k)$ of strings. We use the following encoding

$$(x_1,\ldots,x_k) \equiv x_1 \cdot \ldots \cdot x_k \equiv \ell(x_1)0x_1\ell(x_2)0x_2\cdots\ell(x_{k-1})0x_{k-1}x_k, \tag{12}$$

where $\ell(x_i) = 1^{|x_i|}$ is a string of 1s of length $|x_i|$. Using this encoding, each input x_i adds $2|x_i|+1$ bits to the input string, except the last one, which only adds $|x_i|$; that is

$$|(x_1,\ldots,x_k)| = \sum_{i=1}^{k-1} 2|x_i| + |x_k| + (k-1). \tag{13}$$

Multiple quantum inputs/outputs can be defined by encoding the tuple of sizes (number of qubits) of the different systems in the same way described above.

3.1.2 Symbolic notation for dcq-TM computations

Thus, an input on the dcq-TM of the form $(x;\rho_{\langle 1,n\rangle})$, specifies the following representation of the tapes,

$$\begin{aligned} [\text{Classical tape}] &= [\square]_{(-\infty,0)}\ [x]\ [\square]\ [n]\ [\square]_{(|x|+|n|+1,\infty)} \\ \rho_{\text{ Quantum Tape}} &= |\mathbf{0}\rangle\langle\mathbf{0}|_{(-\infty,0)} \otimes \rho_{\langle 1,n\rangle} \otimes |\mathbf{0}\rangle\langle\mathbf{0}|_{(n+1,\infty)}, \end{aligned} \tag{14}$$

where $|\mathbf{0}\rangle\langle\mathbf{0}|_{(a,b)}$ denotes $|0\rangle\langle 0|_a \otimes |0\rangle\langle 0|_{a+1} \otimes \ldots \otimes |0\rangle\langle 0|_b$. If a specific dcq-TM machine, T, on input $(x;\rho_{\text{in}})$ produces output $(y;\rho_{\text{out}})$ we represent the computation symbolically as,

$$\mathsf{T}(x;\rho_{\text{in}}) = (y;\rho_{\text{out}}), \tag{15}$$

for the overall output, and

$$\mathsf{T}^{(C)}(x;\rho_{\text{in}}) = y \qquad \mathsf{T}^{(Q)}(x;\rho_{\text{in}}) = \rho_{\text{out}}, \tag{16}$$

for the classical and quantum outputs. To represent empty inputs/outputs we use the symbols □ for the classical tape, and ϵ for the quantum one. Note that ϵ is used to denote quantum input/outputs of size

zero. We call the reader's attention to the fact that ϵ is not a density operator acting on a Hilbert space, since a zero size input/output does not refer to any system in the quantum tape; we use it to refer to the cases where there may not be a quantum input/output.

The computability properties of dcq-TM may be dependent on the chosen set $\mathbb{U}$ of basic quantum operations, namely the states that can be prepared exactly by the machine. On this work, we will focus on dcq-TM based on the universal set $\mathbb{U} = \{\mathbb{1}, \mathsf{H}, \mathsf{S}, \frac{\pi}{8}, \mathsf{Sw}, \mathsf{c\text{-}Not}\}$ commonly used in the circuit model for quantum computation. This set is important because it is associated with the capabilities of practical quantum computers. It is also important to remark that the transition function of a dcq-TM depends only on the contents of the classical tape, this is the reason the model is denoted as *deterministic control*. In [16], the universality of this model was proven and, moreover, it was also shown that the dcq-TM model could simulate a special type of quantum circuit called a *seesaw circuit*. A *polynomial translatability theorem* was proven which states that any general quantum circuit can be simulated by one such seesaw circuit with a polynomially bounded overhead on the number of gates, showing that the dcq-TM can solve in a bounded polynomial-time the same problems that the regular quantum circuit model can, therefore, the BQP complexity class coincides with the analogous dcBQP class defined for the dcq-TM model. We also know that there is a universal dcq-TM, T_U, that can simulate every other dcq-TM on any arbitrary input. As such, without loss of generality,

$$\mathsf{T}_\mathsf{U}(x;\rho_{\text{in}}) = (\square;\rho_{\text{out}}) \Leftrightarrow \mathcal{C}_x^*(\rho_{\text{in}}) = \rho_{\text{out}}. \tag{17}$$

The previous expression represents symbolically the equivalence between the computation taken on a specific input in the dcq-TM model, and it is implementation of the *seesaw circuit* $\mathcal{C}_x^*$ that gives the same output. From now on, whenever we refer to a universal dcq-TM, we mean a universal dcq-TM machine fulfilling the s-m-n property.

Sometimes, we may want to refer to an instance of a universal dcq-TM U, simulating some other dcq-TM T. From the s-m-n property, we know that there is a string $t \in \{0,1\}^*$ such that, for any input $(x;\rho_{\text{in}})$ we have that $\mathsf{U}\big((t,x);\rho_{\text{in}}\big) = \mathsf{T}(x;\rho_{\text{in}})$. In these cases we say that the machine U *runs the program* t *on input* $(x;\rho_{in})$, denoted symbolically by:

$$\mathsf{U}_t(x;\rho_{\text{in}}) = \mathsf{U}\big((t,x);\rho_{\text{in}}\big) = \mathsf{T}(x;\rho_{\text{in}}). \tag{18}$$

3.2 Classical quantum state encoding and simulators

In this work we are interested in comparing the complexity aspects of two related, but fundamentally different types of objects. The first are the quantum states themselves, understood as density operators acting on a given Hilbert space, encoded in the physical states of the cells of the quantum tape, and are denoted by standard Greek letters ρ, σ, etc. The second are the *classical representations* of the quantum states, which are understood as arrays of complex numbers describing the respective density matrices in the computational basis, encoded by strings of symbols in the classical alphabet $\mathbb{A}$ according to an appropriate rule (as described in Section 2.2), and denoted by bolded Greek letters $\boldsymbol{\rho}, \boldsymbol{\sigma}$, etc. In the context of inputs and outputs of dcq-TMs, a statement of the form

$$\mathsf{T}(\boldsymbol{\rho}_{\text{in}};\epsilon) = (\boldsymbol{\rho}_{\text{out}};\epsilon), \tag{19}$$

is interpreted as follows: "Upon classical input of any string encoding the computational-basis matrix associated to the state ρ_{in} and no quantum input, the dcq-TM T outputs a string encoding the respective matrix for the state ρ_{out} without a quantum output". Finally, for simplicity, following the relation between the states and their matrix representations, for every function f of quantum states $\rho, \sigma, \dots$, we define $f(\boldsymbol{\rho}, \boldsymbol{\sigma}, \dots) \equiv f(\rho, \sigma, \dots)$.

Before moving on to discuss directly dcq-computable and (plain) dcq-computable quantum states, we establish a notation for the classical simulators of these machines. For any dcq-TM T, consider a "classical" machine $\tilde{\mathsf{T}}$ (a dcq-TM that does not interact with the quantum tape and whose quantum output is ϵ for any input), defined constructively from T, which appropriately approximates the matrix operations associated with the logic gates applied during the execution of T such that

$$\mathsf{T}(x;\rho_{\text{in}}) = (y;\rho_{\text{out}}) \Leftrightarrow \tilde{\mathsf{T}}\big((x,\boldsymbol{\rho}_{\text{in}});\epsilon\big) = \big((y,\boldsymbol{\rho}_{\text{out}});\epsilon\big), \tag{20}$$

whenever the respective string encodings $\boldsymbol{\rho}_{\text{in}}$ and $\boldsymbol{\rho}_{\text{out}}$ exist (that is, whenever the matrices $\boldsymbol{\rho}_{\text{in}}$ and $\boldsymbol{\rho}_{\text{out}}$ are computable).

4 Computability of quantum states

4.1 Directly computable states

Definition 4.1. *A quantum state ρ is directly dcq-computable whenever it can be outputted by a dcq-TM, provided the machine started the computation with no auxiliary input i.e.,*

$$\exists \mathsf{T} : \mathsf{T}(\square;\epsilon) = (\square;\rho). \tag{21}$$

The set of directly dcq-computable states is denoted by $\mathsf{DCOMP_q}$.

Denote by $\mathsf{T}(x;\sigma)\downarrow_t$ a particular (defined constructively from T and t) dcq-TM that simulates $\mathsf{T}(x;\sigma)$, except that if at the step t it has not halted, if forces it to halt, then outputs the output of the simulated machine (according to the rules specified in Section 3). $\mathsf{T}^{(C)}(x;\sigma)\downarrow_t$ and $\mathsf{T}^{(Q)}(x;\sigma)\downarrow_t$ are defined analogously. Consider now Algorithm 1 which, given a universal dcq-TM U, uses Cantor's zig-zag method to assign to each non-negative integer s a different directly dcq-computable state associated with running U for some combination of input and number of steps, which is different for every value of s. Because the output of a dcq-TM is well defined for any state of its tapes, each value of s is uniquely associated with a directly dcq-computable state through Algorithm 1. Additionally, because every directly dcq-computable state ρ is associated with U running *some* program for *some* number of steps, we know that there exists a value of s for which Algorithm 1 outputs ρ. This means that the set of states that can be returned by Algorithm 1 coincides set of directly dcq-computable states. For a fixed reference universal dcq-TM, the quantum output of Algorithm 1 defines a surjective function $\Pi_1(s)$ from the set of non-negative integers to $\mathsf{DCOMP_q}$, that is, an enumeration of it.

Parameters: Universal dcq-TM U ;
Input: Integer $s \geq 0$;
Set n equal to the smallest natural number such that $s < (n+1)(n+2)/2$;
if $n \leq 1$ **then**
  $m = s - n$;
else
  $m = s \mod (n(n+1)/2)$;
end
$\sigma = \mathsf{U}^{(Q)}(n-m;\epsilon)\downarrow_m$;
return $(\square;\sigma)$;

Algorithm 1: directly dcq-computable state generator algorithm.

4.2 Computable states

Definition 4.2. *A quantum state ρ is dcq-computable if there exists a dcq-TM T and an infinite sequence $\{\sigma_i\}_i$ of directly dcq-computable states such that:*

1. $\mathsf{T}(k;\epsilon) = (\square;\sigma_k)$*, i.e., upon classical input k and no quantum input, the machine T outputs the state σ_k.*
2. $D(\rho,\sigma_k) \leq 1/k$*, where $D(\cdot)$ denotes the trace distance of operators acting on the respective Hilbert space.*

We denote the set of all dcq-computable quantum states by $\mathsf{COMP_q}$.

Consider the set $\mathsf{COMP_c}$ of quantum states for which their density matrix representation, when written in the computational basis, has computable complex numbers as components. This set is relevant because it represents the set of quantum states that can be written in a classical computer. Let us state a pair of algorithms that will be useful in finding relationships between quantum states and their representation in the context of the dcq-TM model. Since our machine model is controlled by classical programs, we would expect that $\mathsf{COMP_q}$, the set of computable states as defined by Def. 4.2 and the set $\mathsf{COMP_c}$ coincide. This is indeed the case as shown below.

Theorem 4.1. $\mathsf{COMP_q} = \mathsf{COMP_c}$

Proof.
($\mathsf{COMP_q} \subseteq \mathsf{COMP_c}$):
Let $\rho \in \mathsf{COMP_q}$, and $\{\sigma_i\}_i$ a sequence which satisfies Def. 4.2. From the definition of $\mathsf{COMP_c}$, to show that $\rho \in \mathsf{COMP_c}$ we must show that $\boldsymbol{\rho}$ is computable. Note that the set $\mathbb{U}$ consists on matrices with computable components. It follows that the associated sequence of matrices $\{\boldsymbol{\sigma}_i\}_i$ is computable since, for any dcq-TM T such that $\mathsf{T}(k;\epsilon) = (\square;\sigma_k)$, the classical machine $\tilde{\mathsf{T}}$ simulates the evolution of the initial state though the quantum circuit specified by T and computes the components $\boldsymbol{\sigma}_n$. By recalling that the limit of any classically computable sequence is computable, it follows that $\lim_{n\to\infty} \boldsymbol{\sigma}_n = \boldsymbol{\rho}$ is computable and thus, $\rho \in \mathsf{COMP_c}$.

($\mathsf{COMP_c} \subseteq \mathsf{COMP_q}$):
Let $\boldsymbol{\rho}$ be a density matrix with computable components. From the definition of computability of real numbers, this means that there exists a set $\left\{f^k_{rs}\right\}_{r,s,k}$, of computable functions over $\mathbb{N}$ such that the sequence of matrices $\boldsymbol{\sigma}_k = (g_{rs}(k))$, where $g_{rs}(k) = f^0_{rs}(k) + i f^1_{rs}(k)$, satisfies

$$D(\boldsymbol{\sigma}_k, \boldsymbol{\rho}) \leq \frac{1}{k}. \tag{22}$$

Parameters: Universal dcq-TM U ;
Input: Integer k, $2^n \times 2^n$ density matrix $\boldsymbol{\sigma}$;
$s \leftarrow 0$;
while $\boldsymbol{\sigma}' = \tilde{\Pi}_1(s)$ *is not an* $2^n \times 2^n$ *matrix such that* $D(\boldsymbol{\sigma}', \boldsymbol{\sigma}) \leq 1/(2k)$ **do**
  $s \leftarrow s + 1$;
end
$\sigma' \leftarrow \Pi_1(s)$;
return $(s;\sigma')$

Algorithm 2: Given a density matrix $\boldsymbol{\sigma}$ and integer k, finds a directly dcq-computable state that is closer to σ than $1/(2k)$.

We want to use $\{\boldsymbol{\sigma}_k\}$ to construct a (computable) sequence $\{\sigma'_k\}$ of states associated with directly dcq-computable states that satisfy Def. 4.2. Consider the following construction for σ'_k: Fix a universal dcq-TM U and run Algorithm 2 with input $((k, \boldsymbol{\sigma}_{2k});\epsilon)$ and define σ'_k as its quantum output. We know that Algorithm 2 will always find such state because the set of gates that a dcq-TM has access is universal, and hence the set of directly dcq-computable states is dense in the set of quantum states. To prove that $\rho \in \mathsf{COMP_q}$ it only remains to show that the sequence $\{\sigma_i\}$ satisfies Def. 4.2(2). Indeed, we have that

$$D(\sigma'_k, \rho) \leq D(\sigma'_k, \sigma_{2k}) + D(\sigma_{2k}, \rho) \leq \frac{1}{2k} + \frac{1}{2k} = \frac{1}{k}. \tag{23}$$

□

5 Algorithmic complexity in the dcq-TM model

Having defined the features of the dcq-TM, we proceed to introduce two (non-equivalent) notions of the concept of Kolmogorov complexity. The first one is a generalization of the one proposed in [16] for general directly dcq-computable states:

Definition 5.1. *Let ρ be a directly dcq-computable quantum state and T be a dcq-TM. The basic algorithmic complexity of ρ is*

$$\mathrm{K}^d_{\mathsf{T}}(\rho) = \min_p \left\{ |p| : \mathsf{T}^{(Q)}(p;\epsilon) = \rho \right\}. \tag{24}$$

This definition is limited in that it is only applicable to directly dcq-computable states (denoted by the superscript d), which are dependent on the specific chosen universal set of gates $\mathbb{U}$ in Def. 3.1. We are interested in extending the concept of algorithmic complexity for dcq-computable states.

Definition 5.2. *Let ρ be a quantum state and T be a dcq-TM. The algorithmic complexity of ρ is given by*

$$\mathrm{K}_{\mathsf{T}}(\rho) = \inf\left\{|p| : \forall k \in \mathbb{N}\ D\left(\mathsf{T}^{(Q)}\left((p,k);\epsilon\right),\rho\right) \leq \frac{1}{k}\right\} \tag{25}$$

Definition 5.3. *Let x be a binary string, ρ, σ be two quantum states, and T be a dcq-TM. The conditional algorithmic complexity of ρ given the pair $(x;\sigma)$ is*

$$\mathrm{K}_{\mathsf{T}}(\rho \mid x,\sigma) = \inf\left\{|p| : \forall k \in \mathbb{N}\ D\Big(\mathsf{T}^{(Q)}\left((p,x,k);\sigma\right),\rho\Big) \leq \frac{1}{k}\right\}. \tag{26}$$

In other words, we define the complexity of a quantum state as the infimum length of the *classical description* of the sequence that converges to such a state. Note that, because Equation (25) considers the infimum of the set, states that are not dcq-computable are assigned infinite complexity. This definition is analogous to the standard classical complexity for real numbers (see Def. 2.3) and it is machine independent in the following sense:

Property 5.1. *Let T and T' be universal dcq-TM, then for any string x and quantum states ρ, σ it holds that*

$$\mathrm{K}_{\mathsf{T}}(\rho \mid x;\sigma) = \mathrm{K}_{\mathsf{T}'}(\rho \mid x;\sigma) + O(1) \tag{27}$$

Proof. Let s be a map that translates T'-programs to T-programs and $c \in \mathbb{N}$ be such that $|s(p)| \leq |p| + c$ for every $p \in \{0,1\}^*$. Then,

$$\begin{aligned}
\mathrm{K}_{\mathsf{T}'}(\rho \mid x;\sigma) + c &= \inf\left\{|p| : \forall k \in \mathbb{N}\ D\left(\mathsf{T}'^{(Q)}\left((p,x,k);\sigma\right),\rho\right) \leq \frac{1}{k}\right\} + c \\
&= \inf\left\{|p| : \forall k \in \mathbb{N}\ D\left(\mathsf{T}^{(Q)}\left((s(p),x,k);\sigma\right),\rho\right) \leq \frac{1}{k}\right\} + c \\
&\geq \inf\left\{|s(p)| : \forall k \in \mathbb{N}\ D\left(\mathsf{T}^{(Q)}\left((s(p),x,k);\sigma\right),\rho\right) \leq \frac{1}{k}\right\} \\
&\geq \inf\left\{|p| : \forall k \in \mathbb{N}\ D\left(\mathsf{T}^{(Q)}\left((p,x,k);\sigma\right),\rho\right) \leq \frac{1}{k}\right\} \\
&= \mathrm{K}_{\mathsf{T}}(\rho \mid x;\sigma).
\end{aligned} \tag{28}$$

By a symmetric argument, since T' is also assumed to enjoy the *s-m-n* property, there is $c' \in \mathbb{N}$ such that:

$$\mathrm{K}_{\mathsf{T}}(\rho \mid x;\sigma) + c' \geq \mathrm{K}_{\mathsf{T}'}(\rho \mid x;\sigma). \tag{29}$$

□

Remark. *From Property 5.1 we can talk about the machine independent algorithmic complexity* K *(without subscript) and consider complexities that are equal up to an additive constant to be equivalent.*

We can turn our attention now into exploring the relationship between the complexity of a quantum state and its respective classical representation. As the following theorem states, they turn out to be same.

Theorem 5.2. *For every dcq-TM T and quantum state ρ it holds that*

$$\mathrm{K}(\boldsymbol{\rho}) = \mathrm{K}(\rho). \tag{30}$$

Proof. The result follows directly from the algorithms in the proof of Theorem 4.1, with the constant term being the length of each algorithm's specification. □

An immediate implication of Theorem 5.2 is that, since the components of a density matrix can have arbitrarily high complexity even for single qubit quantum systems, there is no upper bound for the complexity of an n-qubit quantum state. This is in contrast to the complexity of n-bit strings, which is upper bounded by n.[3]

[3]This is also a significant way in which this definition differs from Vitányi's approach, which is also based on classical descriptions, and for which the complexity of an n-bit pure state is upper bounded by $2n$ (see Theorem 3 from [34])

Does Theorem 5.2 mean that the information value of the state of a quantum system exactly coincides with that of the numbers in its density matrix? The reason the two quantities in Equation (30) coincide is because we are quantifying the complexity of a quantum system using a classical string (the description of the program). Note that having access to the description of a program that computes a given quantum state is not the same as having a copy of the quantum state itself. To get a little better insight about how these two types of objects differ we need to consider their use in the role of a *resource*.

5.1 Conditional complexity and state cloning

Let us continue by evidencing a consequence of the no-cloning theorem. Since the machine receives a specification of the size of its quantum input, we know that for any state ρ it holds that $\mathrm{K}(\rho\,|\,\square;\rho) = 0$, as the program that outputs the quantum input has constant length. On the other hand $\mathrm{K}(\rho\otimes\rho\,|\,\square;\rho) \neq 0$, as there is not a single machine that can copy arbitrary quantum states. In the worst case scenario, such machine would construct the second copy of ρ from scratch, which means that the complexity of duplicating a physical state is upper bounded by the complexity of the state itself:

$$\mathrm{K}(\rho\otimes\rho\,|\,\square;\rho) \leq \mathrm{K}(\rho). \tag{31}$$

This is in contrast to the case in which one would want to copy a classical input, in which case it is always possible to output twice the input. If one had *a priori* access to the classical description of the state, it is possible to run Algorithm 2 repeatedly to compute several copies of ρ. Therefore, copying a quantum state given its associated density matrix $\boldsymbol{\rho}$ has constant complexity:

$$\mathrm{K}(\rho\otimes\rho\,|\,\boldsymbol{\rho};\epsilon) = 0. \tag{32}$$

This exercise suggests that the classical description of a quantum state contains "more" descriptive information that the physical state itself, and thus we can conclude that

Property 5.3. *For any dcq-computable quantum states ρ,σ it holds that*

$$\mathrm{K}(\rho\,|\,\boldsymbol{\sigma};\epsilon) \leq \mathrm{K}(\rho\,|\,\square;\sigma). \tag{33}$$

We now turn our attention to answering if, for arbitrary quantum states, the complexity of copying the state is the same as computing it. In other words, we want to know if the converse of Equation (31) holds. Intuitively, this would mean that any machine that computes the duplicate of a state ρ would necessarily have the descriptive information of that state somehow encoded within it. In order to tackle this question, let us first state a couple of lemmas, the proofs of which can be found in the Appendix. Recall that the fidelity between two quantum states ρ and σ is defined as

$$F(\rho,\sigma) = \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2. \tag{34}$$

Lemma 5.4. *Let $T:\mathcal{D}(\mathcal{H})\to\mathcal{D}(\mathcal{H}^{\otimes 2})$ be a CPTP map, i.e., a completely positive and trace-preserving map. Define the* copying error $\mathtt{CE}_T$ *of the transformation T on the state $\rho\in\mathcal{D}(\mathcal{H})$ as*

$$\mathtt{CE}_T(\rho) = \sqrt{1-F(\rho^{\otimes 2},T(\rho))}. \tag{35}$$

For any $\rho_1,\rho_2\in\mathcal{D}(\mathcal{H})$ such that $\mathtt{CE}_T(\rho_1)+\mathtt{CE}_T(\rho_2)\leq\varepsilon\leq\frac{1}{4}$ it holds that

$$F(\rho_1,\rho_2)\leq\frac{1}{2}-\sqrt{\frac{1}{4}-\varepsilon}\quad\text{or}\quad F(\rho_1,\rho_2)\geq\frac{1}{2}+\sqrt{\frac{1}{4}-\varepsilon}. \tag{36}$$

Lemma 5.5. *Let $\mathcal{H}$ be an N-dimensional Hilbert space and $A=\{\rho_i\in\mathcal{D}(\mathcal{H})\}$ be a set of density operators such that for $i\neq j$:*

$$F(\rho_i,\rho_j)\leq\delta_0(N), \tag{37}$$

with

$$\delta_0(N) = \frac{1}{N^2}\left(\frac{1-\sqrt{1-\frac{1}{N}}}{1+2^{\left(\frac{3^N+1}{2}\right)}}\right)^4. \tag{38}$$

Then, A cannot have more than N elements.

Lemma 5.6. *For any CPTP map* $T : \mathcal{D}(\mathcal{H}) \to \mathcal{D}\left(\mathcal{H}^{\otimes 2}\right)$*:*

1. $D\left(T(\rho), \rho^{\otimes 2}\right) \leq 1/k \implies \mathtt{CE}_T(\rho) \leq \sqrt{2/k}$.
2. $D\left(T(\rho), \rho^{\otimes 2}\right) \leq 3D(\rho, \rho') + D\left(T(\rho'), \rho'^{\otimes 2}\right)$.

Theorem 5.7. *For any n-qubit quantum state ρ it holds that*

$$\mathrm{K}(\rho) \leq \mathrm{K}(\rho \otimes \rho \,|\, \square; \rho) + 2n + 2\log(n). \tag{39}$$

Proof. To prove the result, we construct a family of algorithms (See Algorithm. 3) with three parameters: a program index t, a qubit number n, and a list index m. Then show that, given an n-qubit state ρ, there exist $t, m \in \mathbb{N}$ satisfying:

$$|t| \leq K(\rho \otimes \rho | \square; \rho) + c; \qquad |m| \leq n, \tag{40}$$

where c is a constant independent of n, such that the program specified by Algorithm 3 with the parameters (t, n, m) computes ρ. We can see Algorithm 3 as a program that computes ρ from ϵ using a program that computes $\rho \otimes \rho$ from ρ as a subroutine.

Parameters: Universal dcq-TM U. Integers t, n, m;
Input: Integer k;

(List preparation phase)
$L \leftarrow (); s \leftarrow 0; \varepsilon_0 \leftarrow \frac{1}{4}\delta_0(2^n)(1 - \delta_0(2^n)); k_0 \leftarrow \left(\frac{4}{\varepsilon_0}\right)^2$;
while *The size of the list* $|L| \leq m$ **do**
  $\boldsymbol{\sigma} \leftarrow \tilde{\Pi}_1(s)$;
  if $\boldsymbol{\sigma}$ *is a* $2^n \times 2^n$ *matrix such that* $\mathtt{CE}_{T_{t,k_0}}(\boldsymbol{\sigma}) \leq \frac{2}{\sqrt{k_0}}$ & $F(\boldsymbol{\sigma}, \boldsymbol{\sigma}') \leq \delta_0$ *for all* $\boldsymbol{\sigma}' \in L$ **then**
    Add $\boldsymbol{\sigma}$ to L;
  end
  $s \leftarrow s + 1$;
end
$\boldsymbol{\rho}_0 \leftarrow L[m]$ (the m-th element of the list L);

(Output preparation phase)
$s \leftarrow 0; \varepsilon_{\text{out}} \leftarrow \min\left\{\delta_0(2^n)(1 - \delta_0(2^n)), \frac{1}{k^2}\left(1 - \frac{1}{k^2}\right)\right\}; k_{\text{out}} \leftarrow \left(\frac{4}{\varepsilon_{\text{out}}}\right)^2$;
while *no matrix* $\boldsymbol{\rho}_{\text{out}}$ *is found* **do**
  $\boldsymbol{\sigma} \leftarrow \tilde{\Pi}_1(s)$;
  if $\boldsymbol{\sigma}$ *is a* $2^n \times 2^n$ *matrix such that* $\mathtt{CE}_{T_{t,k_{\text{out}}}}(\boldsymbol{\sigma}) \leq \frac{2}{\sqrt{k_{\text{out}}}}$ & $F(\boldsymbol{\sigma}, \boldsymbol{\rho}_0) > \frac{16}{\sqrt{k_0}}$ **then**
    $\boldsymbol{\rho}_{\text{out}} \leftarrow \boldsymbol{\sigma}$;
  else
    $s \leftarrow s + 1$;
  end
end
return $(\square; \rho_{\text{out}} = \Pi_1(s))$;

Algorithm 3: Approximation of an n-qubit state given a subroutine that is known to be able to duplicate it and a specification index $m \leq \log(n)$.

Let U be a reference universal dcq-TM. For any $t, k \in \mathbb{N}$ define the family of CPTP maps $\left\{T_{t,k} : \mathbb{C}^{2^n} \to \mathbb{C}^{2^{2n}}\right\}$ as:

$$T_{t,k}(\sigma) = \mathsf{U}_t^Q(k; \sigma). \tag{41}$$

That is, $T_{t,k}$ represents the transformation (understood as going from the quantum input space to the quantum output space) applied by the machine U running the program t on classical input k.

Let us now gain a little intuition on the two phases of Algorithm 3. The list preparation phase goes through the set of directly dcq-computable states looking for states that are "close" to being copied by T_{t,k_0}. Whenever it finds one such state, it adds it to the list L if it is "far" from all the other states

already in the list. The phase ends when the m-th element is added to the list this way and assigns $\boldsymbol{\rho}_0$ to be equal to the last state added. On the other hand, the output preparation phase goes again through the set of directly dcq-computable states, now looking for a state that is both close to the state ρ_0 obtained from the first phase and "close enough" to being copied by $T_{t,k_{\text{out}}}$. To show that the proposed algorithm satisfies the desired properties, we break down the proof into two parts. We want to show that:

i) If U_t computes $\rho \otimes \rho$ from ρ and for large enough m, the list preparation phase will always add a matrix $\boldsymbol{\rho}_0$ for which
$$F(\boldsymbol{\rho}_0, \boldsymbol{\rho}) \geq 1 - \delta_0(2^n), \tag{42}$$
and such matrix will be found before the list exceeds 2^n elements

ii) On input k, if $\boldsymbol{\rho}_0$ satisfies Equation (42) then the output state $\rho_{\text{out}}(k)$ will satisfy
$$D(\rho_{\text{out}}(k), \rho) \leq \frac{1}{k}. \tag{43}$$

Once these properties are proven, the bound in Equation (39) follows by noting that the minimal size program t' that computes $\rho \otimes \rho$ from ρ can be specified by using $|t'| = K_{\mathsf{U}}(\rho \otimes \rho | \square; \rho) \leq K(\rho \otimes \rho | \square; \rho) + c$ bits, and the values of $m \leq 2^n$ and n can be specified with n and $\log(n)$ bits, respectively.

- Proof of statement i)

Note that whenever a matrix $\boldsymbol{\sigma}$ gets added to L it satisfies by Lemma 5.6(1) and the definition of dcq-computability
$$\mathtt{CE}_{T_{t,k_0}}(\boldsymbol{\sigma}) + \mathtt{CE}_{T_{t,k_0}}(\boldsymbol{\rho}) \leq \frac{2}{\sqrt{k_0}} + \sqrt{\frac{2}{k_0}} \leq \frac{4}{\sqrt{k_0}}, \tag{44}$$
invoking Lemma 5.4 we get that
$$F(\boldsymbol{\sigma}, \boldsymbol{\rho}) \leq \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_0}}} \quad \text{or} \quad F(\boldsymbol{\sigma}, \boldsymbol{\rho}) \geq \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_0}}}, \tag{45}$$
and substituting k_0 in terms of $\delta_0(2^n)$
$$F(\boldsymbol{\sigma}, \boldsymbol{\rho}) \leq \frac{\delta_0}{4} \quad \text{or} \quad F(\boldsymbol{\sigma}, \boldsymbol{\rho}) \geq 1 - \frac{\delta_0}{4}. \tag{46}$$
Depending on the transformation T_{t,k_0}, the loop inside the **while** statement of the list preparation phase will add some amount of elements to the list before there are no more states that satisfy the stated conditions. Let $l_{\max}$ be the maximum amount of elements that can be added to L in this way. From Lemma 5.5 we know that the $l_{\max} \leq 2^n$ because the largest size of a set $\{\boldsymbol{\sigma}_i\}$ that satisfies $F(\boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j) \leq \delta_0(\mathbb{C}^{2^n})$ for $i \neq j$ is equal to $\dim(\mathbb{C}^{2^n}) = 2^n$. Recall now that the set of n-qubit directly dcq-computable states is dense in the set of n-qubit quantum states and the fact that the function $\tilde{\Pi}_1(s)$ will eventually output every directly dcq-computable state. Let s' be the smallest integer such that
$$D(\boldsymbol{\rho}' = \tilde{\Pi}_1(s'), \boldsymbol{\rho}) \leq \frac{1}{3k_0}, \tag{47}$$
using Lemma 5.6(2) and the fact that U_t computes $\rho \otimes \rho$ from ρ
$$D\left(T_{t,k_0}(\boldsymbol{\rho}'), \boldsymbol{\rho}'^{\otimes 2}\right) \leq 3D(\boldsymbol{\rho}', \boldsymbol{\rho}) + D\left(T_{t,k_0}(\boldsymbol{\rho}), \boldsymbol{\rho}^{\otimes 2}\right) \tag{48}$$
$$\leq \frac{2}{k_0}, \tag{49}$$
and from Lemma 5.6(1) we get
$$\mathtt{CE}_{T_{t,k_0}}(\boldsymbol{\rho}') \leq \frac{2}{\sqrt{k_0}}. \tag{50}$$

Suppose the loop reaches s' without finding $\boldsymbol{\sigma}$ satisfying Equation (42). Then, as shown in Equation (50), at $s = s'$ the matrix $\boldsymbol{\rho}' = \tilde{\Pi}_1(s')$ will satisfy the first condition in the check to be added to L. Additionally, for every $\boldsymbol{\sigma}$ already in L at that point, it will hold that (see Lemma 1 from [27])

$$\sqrt{1 - F(\boldsymbol{\rho}', \boldsymbol{\sigma})} \geq |\underbrace{F(\boldsymbol{\rho}', \boldsymbol{\rho})}_{\geq 1 - \frac{\delta_0}{4}} - \underbrace{F(\boldsymbol{\rho}, \boldsymbol{\sigma})}_{\leq \frac{\delta_0}{4}}| \geq 1 - \frac{\delta_0}{2}, \tag{51}$$

and hence

$$F(\boldsymbol{\rho}', \boldsymbol{\sigma}) \leq \delta_0 \left(1 - \frac{\delta_0}{4}\right) \leq \delta_0, \tag{52}$$

which satisfies the second condition to be added to L. Therefore, for large enough m, the list L will always have an element that satisfies Equation (42). Particularly, there is a value of $m \leq l_{\max} \leq 2^n$ for which the cycle ends when such element is added and sets it to be $\boldsymbol{\rho}_0$.

- Proof of statement ii)

Suppose $\boldsymbol{\rho}_0$ satisfies Equation (42). We proceed to show that any density matrix $\boldsymbol{\sigma}$ that holds both of the conditions in the **if** statement of the output preparation phase must also satisfy

$$D(\boldsymbol{\sigma}, \boldsymbol{\rho}) \leq \frac{1}{k}. \tag{53}$$

Analogously to Equations (44) and (45), whenever a matrix $\boldsymbol{\sigma}$ is found that satisfies $\texttt{CE}_{T_{t,k_{\text{out}}}}(\boldsymbol{\sigma}) \leq 2/\sqrt{k_{\text{out}}}$ we have that

$$\texttt{CE}_{T_{t,k_{\text{out}}}}(\boldsymbol{\sigma}) + \texttt{CE}_{T_{t,k_{\text{out}}}}(\boldsymbol{\rho}) \leq \frac{2}{\sqrt{k_{\text{out}}}} + \sqrt{\frac{2}{k_{\text{out}}}} \leq \frac{4}{\sqrt{k_{\text{out}}}}, \tag{54}$$

and

$$F(\boldsymbol{\sigma}, \boldsymbol{\rho}) \leq \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_{\text{out}}}}} \quad \text{or} \quad F(\boldsymbol{\sigma}, \boldsymbol{\rho}) \geq \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_{\text{out}}}}}. \tag{55}$$

In other words, such $\boldsymbol{\sigma}$ is either "very close" to $\boldsymbol{\rho}$ or "very far" from it. Consider the case where left part of Eq (55) is true, we have that (noting that $k_{\text{out}} \geq k_0$)

$$\begin{aligned}
\sqrt{1 - F(\boldsymbol{\rho}_0, \boldsymbol{\sigma})} &\geq | \underbrace{F(\boldsymbol{\rho}_0, \boldsymbol{\rho})}_{\geq \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_0}}}} - \underbrace{F(\boldsymbol{\rho}, \boldsymbol{\sigma})}_{\leq \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_{\text{out}}}}}} | \\
&\geq \left| \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_0}}} + \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_{\text{out}}}}} \right| \\
&\geq \sqrt{1 - \frac{16}{\sqrt{k_0}}},
\end{aligned} \tag{56}$$

and therefore

$$F(\boldsymbol{\rho}_0, \boldsymbol{\sigma}) \leq \frac{16}{\sqrt{k_0}}, \tag{57}$$

which means that such $\boldsymbol{\sigma}$ cannot satisfy the second condition of the **if** statement. We conclude that for any $\boldsymbol{\sigma}$ that satisfies both conditions (and hence also for $\boldsymbol{\rho}_{\text{out}}$), it holds that

$$\begin{aligned}
F(\boldsymbol{\sigma}, \boldsymbol{\rho}) &\geq \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{4}{\sqrt{k_{\text{out}}}}} \\
&\geq \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{k^2}\left(1 - \frac{1}{k^2}\right)} \\
&= 1 - \frac{1}{k^2}.
\end{aligned} \tag{58}$$

Finally, translating into trace distance we get

$$D(\boldsymbol{\sigma}, \boldsymbol{\rho}) \leq \sqrt{1 - F(\boldsymbol{\sigma}, \boldsymbol{\rho})} \leq \frac{1}{k} \tag{59}$$

□

One way to interpret Theorem 5.7 along with Equation (32) is that, for the task of describing $\rho \otimes \rho$, the advantage of having prior access to a single copy of ρ as compared to not having any prior resource state is at most $2n + 2\log(n)$ bits (up to an additive constant). This means that for complex quantum states, for which $\mathrm{K}(\rho) \gg n$, having access to a copy of ρ gives effectively *no advantage* for describing $\rho \otimes \rho$. This is in contrast with classical strings, where having a copy of the string gives all the information needed to compute its duplicate, and can be understood as an algorithmic information version of the no-cloning theorem.

6 On the chain rule and complexity of quantum correlations

One feature of quantum multipartite states is that, in general, they contain correlations between their parts. One way to quantify correlations is through the notion of mutual information. In this section we use the dcq-TM model to generalize the chain rule and extend the concept of algorithmic mutual information to quantum states. Let us first define the concept of prefix for quantum states.

Definition 6.1. *(Prefix of a quantum state)*
Let $\rho_{\langle 1,n\rangle} \in \mathcal{D}\left(\mathbb{C}^{2^n}\right)$. For $m \leq n$, we say that $\sigma \in \mathcal{D}(\mathbb{C}^{2^m})$ is a prefix of $\rho_{\langle 1,n\rangle}$ whenever

$$\sigma = \mathrm{Tr}_{(m+1,\ldots,n)}\left[\rho_{\langle 1,n\rangle}\right]. \tag{60}$$

We can make use of quantum prefixes to study the behaviour of K when considering systems split into two (or more) parts. Note that an n-qubit state has n possible prefixes, which are manifestly defined by the layout of the cells in the quantum tape. When we want to stress the number m of qubits of a prefix we will refer to it as the "m-prefix". To make the separation more explicit in some of the following definitions we will denote the larger n-qubit space as $\mathcal{H}_{AB}$ and the m-qubit prefix space as $\mathcal{H}_A$ so we can write a quantum state and its prefix as ρ_{AB} and ρ_A, respectively.

By analogy with the classical algorithmic information theory, it seems natural to begin our exploration of mutual information by considering the straightforward generalization of two (classically) related quantities:

Definition 6.2. *(Algorithmic mutual information 1)*
Let ρ_A and ρ_B be two quantum states. Define the mutual information of ρ_A respect to ρ_B by:

$$\mathrm{I}^{(1)}(\rho_A : \rho_B) = \mathrm{K}(\rho_A) - \mathrm{K}(\rho_A \mid \boldsymbol{\rho}_B). \tag{61}$$

Definition 6.3. *(Algorithmic mutual information 2)*
Let ρ_{AB} be a bipartite quantum state. Define the mutual information between the partial states ρ_A and ρ_B by:

$$\mathrm{I}^{(2)}(\rho_A : \rho_B) = \mathrm{K}(\rho_A) + \mathrm{K}(\rho_B) - \mathrm{K}(\rho_{AB}). \tag{62}$$

Definition 6.2 compares the algorithmic complexity of the state ρ_A assuming the machine starts with the all zero state in its classical and quantum tapes, versus its complexity assuming the machine has the description of the state ρ_B given as a classical input. The choice of giving the classical description instead the state itself will become clearer as we study its potential in the context of correlations. Evidently $\mathrm{K}(\rho_A|\boldsymbol{\rho}_B)$ is upper bounded by $\mathrm{K}(\rho_A)$, and hence $\mathrm{I}^{(1)}$ is non-negative. On the other hand, Def. 6.3 considers the difference between computing the states ρ_A and ρ_B separately as opposed to computing the (potentially correlated) single state ρ_{AB}. The main conceptual difference between $\mathrm{I}^{(1)}$ and $\mathrm{I}^{(2)}$ is whether to consider the states ρ_A and ρ_B to exist independently or as a part of a greater joint state. Under the classical definition of Kolmogorov complexity, the analogous expressions to (61) and (62) for mutual information turn out to be connected through the chain rule as shown by Equation (7). To explore what is the connection between these two quantities, we may ask ourselves how would a quantum

version of the chain rule look like? A straightforward generalization of Equation (7) would be to replace the string xy with a bipartite state ρ_{AB} leading to an expression of the form:

$$\mathrm{K}(\rho_{AB}) \underset{?}{=} \mathrm{K}(\rho_A) + \mathrm{K}(\rho_B \mid \boldsymbol{\rho}_A) + O\big(\log(n\,\mathrm{K}(\rho_{AB}))\big), \tag{63}$$

but as Theorem 6.1 states, the above equality does not hold for arbitrary joint states.

Theorem 6.1. *There exist no constant g such that*

$$\mathrm{K}(\rho_{AB}) \leq \mathrm{K}(\rho_A) + \mathrm{K}(\rho_B \mid \boldsymbol{\rho}_A) + g\,\log(n\,\mathrm{K}(\rho_{AB})) \tag{64}$$

for every natural number n and n-qubit quantum bipartite state ρ_{AB}.

Proof. Let $g \in \mathbb{R}$, $\rho_{AB} \in \mathcal{D}((\mathbb{C}^2)^{\otimes n=n_a+n_b})$ such that $\rho_A \in \mathcal{D}((\mathbb{C}^2)^{\otimes n_a})$ and $\rho_B \in \mathcal{D}((\mathbb{C}^2)^{\otimes n_b})$. Let $\rho_A = \sum_{i=1}^{2^{n_a}} \alpha_i |i\rangle\langle i|$ and $\rho_B = \sum_{j=1}^{2^{n_b}} \beta_i |j\rangle\langle j|$ be the spectral decomposition of the reduced systems ρ_A and ρ_B. Assume that the Inequality (64) holds for some ρ_{AB} for which ρ_A and ρ_B are proper mixed states and some constant g. We show that there exists then a $\tilde{\rho}_{AB}$ for which the inequality does not hold. Hence, there is no g that satisfies Equation (64) for all bipartite mixed states.

Note that since ρ_A and ρ_B are proper mixed states, there exist r, r' and s, s' such that $r \neq r'$, $s \neq s'$, and $\alpha_r, \alpha_{r'}, \beta s, \beta s'$ are all non-zero. Define the set of states

$$\tilde{\mathcal{C}} = \left\{ \tilde{\rho}^{\lambda}_{AB} \mid \tilde{\rho}^{\lambda}_{AB} = \lambda(\rho_A \otimes \rho_B) + (1-\lambda)\sigma_{AB}, \text{ such that } \lambda \in (0,1) \text{ is computable } \right\}, \tag{65}$$

where $\rho_A \otimes \rho_B = \sum_{i,j=1}^{2^{n_a}, 2^{n_b}} \alpha_i \beta_j |i,j\rangle\langle i,j|$, and

$$\sigma_{AB} = \rho_A \otimes \rho_B + \gamma(|r,s\rangle\langle r',s'| + |r',s'\rangle\langle r,s|), \tag{66}$$

where $\gamma = \sqrt{\alpha_r \alpha_{r'} \beta_s \beta_{s'}}$. Note that σ_{AB} and $\rho_A \otimes \rho_B$ are strictly different density operators and, since the partial traces on the second term of σ_{AB} both vanish, we have that for any $0 \leq \lambda \leq 1$

$$\begin{aligned}
\tilde{\rho}^{\lambda}_A &= \mathrm{Tr}_B \tilde{\rho}^{\lambda}_{AB} \\
&= \lambda \mathrm{Tr}_B[\rho_A \otimes \rho_B] + (1-\lambda)\mathrm{Tr}_B \sigma_{AB} \\
&= \lambda \mathrm{Tr}_B[\rho_A \otimes \rho_B] + (1-\lambda)\left(\mathrm{Tr}_B[\rho_A \otimes \rho_B] + \gamma \mathrm{Tr}_B\left[|r,s\rangle\langle r',s'| + |r',s'\rangle\langle r,s|\right]\right) \\
&= \rho_A,
\end{aligned} \tag{67}$$

and similarly $\tilde{\rho}^{\lambda}_B = \rho_B$. Notice also that, since ρ_A and ρ_B are assumed to be dcq-computable, it follows from Theorem 4.1 all of the $\tilde{\rho}^{\lambda}_{AB}$ are dcq-computable quantum states. Let us denote the value $g_0 = \mathrm{K}(\rho_A) + \mathrm{K}(\rho_B \mid \boldsymbol{\rho}_A)$. We can write the inequality (64) as:

$$\mathrm{K}(\rho_{AB}) - g\log(n\,\mathrm{K}(\rho_{AB})) - g_0 \leq 0. \tag{68}$$

Define now the set

$$\mathcal{C}_m = \{\rho'_{AB} \mid \rho'_{AB} \in \mathcal{D}((\mathbb{C}^2)^{\otimes n}), \mathrm{K}(\rho'_{AB}) \leq m\}. \tag{69}$$

Note that for all m, the set $\mathcal{C}_m$ is finite (of size at most 2^m), whereas the set $\tilde{\mathcal{C}}$ is infinite (its size is equal to the set of computable real numbers between 0 an 1), this means that for any m, there exists λ such that $\tilde{\rho}^{\lambda}_{AB} \notin \mathcal{C}_m$ ($K(\tilde{\rho}^{\lambda}_{AB}) \geq m$). In other words, given a bipartite state ρ_{AB} with mixed partial traces, there are an infinite number of states ρ'_{AB} that share its partial traces and have larger quantum Kolmogorov complexity under the dcq-TM model. Additionally, for any n, g, g_0 we have that

$$\lim_{m\to\infty} \big(m - g\log(nm) - g_0\big) = \infty. \tag{70}$$

Therefore, there is an m_0 such that $m' - g\log(nm') - g_0 > 0$ for any $m' \geq m_0$. The result follows by noting that, from the above argument, there exists λ such that $\mathrm{K}(\tilde{\rho}^{\lambda}_{AB}) \geq m_0$, and hence

$$\mathrm{K}(\tilde{\rho}^{\lambda}_{AB}) > \mathrm{K}(\tilde{\rho}^{\lambda}_A) + \mathrm{K}(\tilde{\rho}^{\lambda}_B \mid \tilde{\boldsymbol{\rho}}^{\lambda}_A) + g\,\log(n\,\mathrm{K}(\tilde{\rho}^{\lambda}_{AB})). \tag{71}$$

□

To understand better relation between Def. 6.2 and Def 6.3, we have to consider the role of correlations in the description of bipartite states. In the Shannon approach, the notion of information is based on the knowledge of system's state: the more we are uncertain of the state of the system, the more we increase our knowledge (i.e., information) upon learning that state. Thus, mutual information between two systems is manifestly linked to correlations between them. In the original Kolmogorov approach, the notion of information is based on our ability to compute a given string. Consequently, mutual information between two strings tells us how much knowing one helps in computing the other. When generalizing to the quantum domain, we see that part of this intuition holds for Def. 6.2 as well. In other words, algorithmic mutual information is quantifying the similarity between the descriptions of two states, complexity wise, and says noting about the correlations in their joint state. Indeed, the quantities on the right hand side of Equation (61) do not depend on the joint state ρ_{AB} at all. On the other hand, Def. 6.3 effectively compares the complexity of computing a (generally) correlated state with its individual parts. One problem with Def. 6.3 as a measure of complexity is that, following Theorem 6.1, the quantity is unbounded from below.

At this point we can ask ourselves: is there a version of the chain rule that holds for quantum states? Such expression would help us motivate an alternative definition for algorithmic mutual information that satisfies a form of symmetry, as well as give us insight into the complexity of correlations in quantum systems. Theorem 6.2 provides a relatively straightforward generalization of Equation (7) by introducing a direct dependence on the joint state, as well as requiring a classical description of the prefix (as opposed to a copy of the state).

Theorem 6.2. *Let n, m be integers such that $1 < m \leq n$, and ρ be an n-qubit quantum state. For any m-prefix σ of ρ it holds that*

$$\mathrm{K}(\rho) = \mathrm{K}(\sigma) + \mathrm{K}(\rho|\boldsymbol{\sigma}) + O\left(\log(n\,\mathrm{K}(\rho))\right). \tag{72}$$

Proof. Let U be a reference universal dcq-TM:
($\leq$):
We specify an algorithm to compute ρ given as input two parameters: the specifications t_1, t_2 of two programs that compute σ and $\rho|\boldsymbol{\sigma}$, respectively. The algorithm works as follows: on input k, simulates the action of $\tilde{\mathsf{U}}_{t_1}(2k;\varepsilon)$ to obtain $\boldsymbol{\sigma_k}$, then simulate $\mathsf{U}_{t_2}((2k,\boldsymbol{\sigma_k});\varepsilon)$ to obtain ρ'_k. Let $\rho_k = \mathsf{U}_{t_2}((2k,\boldsymbol{\sigma});\varepsilon)$, using the triangle inequality and the contractive property of the trace distance we get that

$$\begin{aligned} D(\rho,\rho'_k) &\leq D(\rho,\rho_k) + D(\rho_k,\rho'_k) \\ &\leq D(\rho,\rho_k) + D(\boldsymbol{\sigma},\boldsymbol{\sigma_k}) \\ &\leq \frac{1}{2k} + \frac{1}{2k} = \frac{1}{k}. \end{aligned} \tag{73}$$

So the algorithm correctly computes ρ. Specifying t_1, t_2 can be done using $\mathrm{K}(\sigma) + \mathrm{K}(\rho|\boldsymbol{\sigma})$ bits. Since this is a two-input program, a specification of the length of one of the inputs must be added, which is of size at least $\min\{\log(\mathrm{K}(\sigma)), \log(\mathrm{K}(\rho|\boldsymbol{\sigma}))\} \leq \log(\mathrm{K}(\rho)) + \log(m)$. We can conclude then that

$$\mathrm{K}(\rho) \leq \mathrm{K}(\sigma) + \mathrm{K}(\rho|\boldsymbol{\sigma}) + \log(m\,\mathrm{K}(\rho)) \leq \mathrm{K}(\sigma) + \mathrm{K}(\rho|\boldsymbol{\sigma}) + \log(n\,\mathrm{K}(\rho)) \tag{74}$$

($\geq$):
By *reductio ad absurdum*, assume that for every positive real number c there exist an integer $n > 1$, an n-qubit quantum state ρ, and a prefix σ of ρ such that

$$\mathrm{K}(\rho) < \mathrm{K}(\sigma) + \mathrm{K}(\rho|\boldsymbol{\sigma}) - c\log(n\,\mathrm{K}(\rho)). \tag{75}$$

Let (c, ρ, σ) be such that they satisfy Equation (75) with ρ, σ being n and $m \leq n$ qubit states, respectively. Furthermore, let t_ρ be such that U_{t_ρ} computes ρ and $|t_\rho| = \mathrm{K}(\rho) + O(1)$. In the following treatment, it will be useful to have a shorthand notation for the prefixes of the quantum output of programs. Let $\eta(t,x)$ denote the m-prefix of $\mathsf{U}^Q_t(x;\varepsilon)$. For any positive integer s consider now the following sets:

$$P_s = \left\{ t \in \{0, 2^{|t_\rho|}\} \middle| \text{ for all } 1 \leq x \leq s,\ \mathsf{U}_t(x;\varepsilon) \text{ halts with a quantum output of at least } m \text{ qubits} \right\} \tag{76}$$

$$P_s^\eta = \left\{ t \in P_s \middle| \text{ for all } 1 \leq x \leq s,\ D(\eta(t,x),\eta) \leq \frac{1}{x} \right\} \tag{77}$$

Recall that a set A is computably enumerable if there exists a bijective function $\mathcal{E}_A : \{1, \ldots, |A|\} \to A$ and a machine T_A such that $\mathsf{T}_A(i;\varepsilon) = (\mathcal{E}_A(i);\varepsilon)$ for every $i = \{1, \ldots, |A|\}$. To see that the set P_s^η is computably enumerable given $s, m, |t_\rho|$, and $\boldsymbol{\eta}$, consider a machine that classically simulates instances of all programs up to length $|t_\rho|$ running with all inputs $x \in \{1, \ldots, s\}$ in parallel, keeping track of their simulated outputs as they halt, and outputting the i-th program found to satisfy the conditions in Equation (77). Denote the enumeration of P_s^η defined by this machine as $\mathcal{E}_{s,m,|t_\rho|,\boldsymbol{\eta}}$.

We want to use the set P_s^σ to find an upper bound for $\mathrm{K}(\rho|\boldsymbol{\sigma})$. Because we know that $t_\rho \in P_s^\sigma$ for any s, fix $s = s_0$ and let j be the index assigned to t_ρ by $\mathcal{E}_{s_0,m,|t_\rho|,\boldsymbol{\sigma}}$. We can compute $\rho|\boldsymbol{\sigma}$ as follows: consider a machine that on input $(k;\boldsymbol{\sigma})$ outputs the values of $|t_\rho|$, m, and j, then finds the value of $t_\rho = \mathcal{E}_{s_0,m,|t_\rho|,\boldsymbol{\sigma}}(j)$ and outputs the simulated quantum output of $\mathsf{U}_{t_\rho}(k)$. By recalling that $|m| \leq \log(n)$ we get that

$$\mathrm{K}(\rho|\boldsymbol{\sigma}) \leq \log(|P_{s_0}^\sigma|) + 2\log(\mathrm{K}(\rho)) + 2\log(n) = \log(|P_{s_0}^\sigma|) + 2\log(n\,\mathrm{K}(\rho)). \tag{78}$$

Using Equation (75) we can obtain a bound for the size of $P_{s_0}^\sigma$

$$\log(|P_{s_0}^\sigma|) > \mathrm{K}(\rho) - \mathrm{K}(\sigma) + (c-2)\log(n\,\mathrm{K}(\rho)) = \ell \tag{79}$$

which is independent of the value of s_0. To arrive to a contradiction, we will proceed to show that the bound found in Equation (79) is "too big", in the sense that it allows for a very short description of σ. To do this, we will use the fact that we can compute σ by simulating a program that computes ρ and then outputting only the first m qubits of its quantum output. In general, for input k, we do not really need to run a program that computes ρ, instead we can run any program that outputs a state "close enough" to the output of $\mathsf{U}_{t_\rho}(ak)$ for some natural number $a \geq 2$. We want to construct a small set of programs that are candidates to approximate σ in this way for a given k and then specify one of them by giving its index in an enumeration of that set. First, consider a slightly relaxed version of the P_s^η family of sets

$${}_2P_s^\eta = \left\{ t \in P_s | \text{ for all } 1 \leq x \leq s,\ D(\eta(t,x),\eta) \leq \frac{2}{x} \right\}. \tag{80}$$

Clearly, $P_s^\eta \subseteq {}_2P_s^\eta$ for all parameters s, η, and they are computably enumerable in the same way the P_s^η are. Additionally, note that for any $t \in P_s^\sigma$ and $1 \leq x \leq s$ it holds that

$$D(\eta(t,x),\eta(t_\rho,s)) \leq D(\eta(t,x),\sigma) + D(\sigma,\eta(t_\rho,s)) \leq \frac{1}{x} + \frac{1}{s} \leq \frac{2}{x}, \tag{81}$$

and thus, for any s we have that $P_s^\sigma \subseteq {}_2P_s^{\eta(t_\rho,s)}$ and, following Equation (79), we get that

$$|{}_2P_s^{\eta(t_\rho,s)}| > 2^\ell. \tag{82}$$

Define now the sets

$$\Sigma_s = \left\{ t \in P_s |\ |{}_2P_s^{\eta(t,s)}| > 2^\ell \right\}. \tag{83}$$

The set Σ_s is computably enumerable as shown by Algorithm 5 (see Appendix). Because we know that $t_\rho \in \Sigma_s$ can be used to approximate σ (given the value of m), we can think of the $\Sigma_s \subseteq P_s$ as a subset of candidates for programs to compute σ. Unfortunately, it is not straightforward to find a bound for $|\Sigma_s|$. Instead, define Σ_s^* as the set of states that can be outputted by Algorithm 4, which constructs a list of programs in Σ_s that are all at least $\frac{4}{s}$-distant from each other. The set Σ_s^* is computably enumerable by construction and satisfies the following properties:

1. For any distinct $t, t' \in \Sigma_s^*$ it holds that the associated ${}_2P_s^{\eta(t,s)}$, ${}_2P_s^{\eta(t',s)}$ are disjoint.
2. There is at least one $\tau \in \Sigma_s^*$ such that $D(\eta(t_\rho,s),\eta(\tau,s)) \leq \frac{4}{s}$ and therefore

$$D(\sigma,\eta(\tau,s)) \leq D(\sigma,\eta(t_\rho,s)) + D(\eta(t_\rho,s),\eta(\tau,s)) \leq \frac{1}{s} + \frac{4}{s} = \frac{5}{s} \tag{84}$$

Parameters: Universal dcq-TM U, enumeration $\mathcal{E}^{\Sigma}_{s,m,|t_\rho|}$ of the set Σ_s given $s, m, |t_\rho|, \ell$;
Input: Integers $i, s, m, |t_\rho|, \ell$;
$L \leftarrow ()$, $j \leftarrow 1$;
while $|L| < i$ **do**
  $t \leftarrow \mathcal{E}^{\Sigma}_{s,m,|t_\rho|}(j)$;
  if $D(\eta(t,s), \eta(t',s)) > \frac{4}{s}$ *for each* $t' \in L$ **then**
    Add t to L;
  end
  $j \leftarrow j + 1$;
end
return $L[i]$

Algorithm 4: Enumeration of the set Σ^*_s.

This means that we can compute σ as follows. On input k, let j be the index of a program τ that satisfies Equation (84) for the set Σ^*_{5k}. Directly compute the values of $m, |t_\rho|$, and ℓ; then run Algorithm 4 on input $(j, 5k, m, |t_\rho|, \ell)$ to obtain τ, and finally output the m-prefix of $\mathsf{U}^Q_\tau(5k;\epsilon)$, which by Equation (84) is at most $\frac{1}{k}$-distant from σ. The size of the index $j \leq |\Sigma^*_{5k}|$ can be bounded by

$$|\Sigma^*_{5k}| < \frac{2^{\mathrm{K}(\rho)+O(1)}}{2^\ell}, \tag{85}$$

by noting that

$$2^\ell |\Sigma^*_{5k}| = \sum_{t \in \Sigma^*_{5k}} 2^\ell < \sum_{t \in \Sigma^*_{5k}} {}_2P^{\eta(t,5k)}_{5k} \leq |P_s| \leq 2^{\mathrm{K}(\rho)+O(1)}, \tag{86}$$

where we used the fact that the considered sets ${}_2P^{\eta(t,5k)}_{5k}$ are disjoint. Equation (85) in turn gives us a bound for the complexity of σ

$$\begin{aligned} \mathrm{K}(\sigma) &\leq \log(|\Sigma^*_{5k}|) + 2\log(n) + 2\log(\mathrm{K}(\rho)) + 2\log(\ell) \\ &< \mathrm{K}(\rho) - \ell + 2\log(n\,\mathrm{K}(\rho)) + 2\log(\ell), \end{aligned} \tag{87}$$

substituting ℓ from Equation (79) and cancelling terms

$$\begin{aligned} 0 &< 2\log\big(\mathrm{K}(\rho) - \mathrm{K}(\sigma) + (c-2)\log(n\,\mathrm{K}(\rho))\big) - (c-4)\log(n\,\mathrm{K}(\rho)) \\ &\leq 2\log\big(n\,\mathrm{K}(\rho) + c\log(n\,\mathrm{K}(\rho))\big) - (c-4)\log(n\,\mathrm{K}(\rho)) \\ &\leq 2\log\big((c+1)n\,\mathrm{K}(\rho)\big) - (c-4)\log(n\,\mathrm{K}(\rho)) \\ &\leq 2\log(c+1) - (c-6)\log(n\,\mathrm{K}(\rho)) \\ &\leq 2\log(c+1) - (c-6) \qquad (\text{for } c \geq 6 \text{ and } n > 1). \end{aligned} \tag{88}$$

We arrive to the contradiction by noting that the last expression on Equation (88) is negative for large enough c. □

By applying Theorem 6.2 to a bipartite state ρ_{AB} we get our new version of the chain rule

$$\begin{aligned} \mathrm{K}(\rho_{AB}) &= \mathrm{K}(\rho_A) + \mathrm{K}(\rho_{AB}|\boldsymbol{\rho}_A) + O\left(\log(n\,\mathrm{K}(\rho_{AB}))\right) \\ &= \mathrm{K}(\rho_B) + \mathrm{K}(\rho_{AB}|\boldsymbol{\rho}_B) + O\left(\log(n\,\mathrm{K}(\rho_{AB}))\right), \end{aligned} \tag{89}$$

where the second equality comes from the fact that the state ρ_{AB} and the state obtained by swapping the order of the subsystems A and B are related by a $\log(n)$ overhead, which gets absorbed by the last term in Equation (72). Notably, Equation (89) recovers the form of the classical chain rule Equation (7) when the two subsystems are uncorrelated, that is

$$\mathrm{K}(\rho_A \otimes \rho_B) = \mathrm{K}(\rho_A) + \mathrm{K}(\rho_B|\boldsymbol{\rho}_A) + O\left(\log(n\,\mathrm{K}(\rho_A \otimes \rho_B))\right). \tag{90}$$

Equipped by this result, we can define a correlation-accounting version of mutual information by quantifying by "how much" a state deviates from the classical chain rule, we call this quantity *complexity of correlations*:

Definition 6.4. *(Complexity of quantum correlations)*
Let ρ_{AB} be a bipartite quantum state. Define the complexity of correlations in the subsystem A respect to B by:

$$\begin{aligned} C(A:B) &= \mathrm{K}(\rho_{AB}|\boldsymbol{\rho}_B) - \mathrm{K}(\rho_A|\boldsymbol{\rho}_B) \\ &= I^{(1)}(A:B) - I^{(Q)}(A:B), \end{aligned} \tag{91}$$

where

$$I^{(Q)}(A:B) = \mathrm{K}(\rho_A) - \mathrm{K}(\rho_{AB}|\boldsymbol{\rho}_B). \tag{92}$$

Because $\boldsymbol{\rho}_B$ has the information of the dimension of the space $\mathcal{H}_B$, it is straightforward to see that $C(A:B)$ is non-negative (up to a constant). Note that the quantity $I^{(Q)}(A:B)$ should not be interpreted as an information measure because it can be negative. In fact, it is generically negative, reaching its maximum zero value for product states $\rho_A \otimes \rho_B$ (we denote the quantity by $I^{(Q)}$ to highlight its analogous formal mathematical expression to $I^{(1)}$). From Equations (89) and (90) we know that C also satisfies symmetry of information: for an n-qubit state ρ_{AB} it holds that

$$\begin{aligned} C(A:B) - C(B:A) &= \underbrace{\big(I^{(1)}(A:B) - I^{(1)}(B:A)\big)}_{\underbrace{\in O(\log(n\,\mathrm{K}(\rho_A\otimes\rho_B)))}_{\in O(\log(n\,\mathrm{K}(\rho_{AB})))}} - \underbrace{\big(I^{(Q)}(A:B) - I^{(Q)}(B:A)\big)}_{\in O(\log(n\,\mathrm{K}(\rho_{AB})))} \\ &= O\big(\log(n\,\mathrm{K}(\rho_{AB}))\big). \end{aligned} \tag{93}$$

7 Discussion

In this work, we have laid the foundations of a quantum algorithmic complexity theory based on the dcq-TM model. We started by expanding the definition of the machine to allow it to work with mixed states as inputs/outputs and defined the set of computable quantum states. The choice of our machine model is motivated by two of its main properties: a) the set of machines is naturally discrete and shown to be Turing-complete, and b) there is a single, well defined, halting state for every machine. We mention this properties in contrast to quantum-controlled type of machine models whose description may be parameterized by the set of quantum states associated to a given Hilbert space. If the set of machines is described by a continuous set of states, it can lead to super-Turing capabilities.[4] This issue can be solved by allowing only some discrete subset of states (chosen adequately to output approximations of quantum states within a desired accuracy), as described in [6]. Another, more difficult to solve issue, is that machines running quantum programs do not always have a single, well defined, halting state. Instead, due to the unitary evolution principle of quantum mechanics, they are generally in a superposition of halting and working states, which makes defining computability difficult. This issue has been openly discussed (see, for instance [5, 15, 19, 25]). We make the observation that, in practical scenarios, for adequately modeling real-life quantum computers, we see ourselves as classical users interacting with the quantum machine, which is in line with the dcq-TM model.

Another advantage of the dcq-TM model is that, by having explicitly separated classical and quantum tapes, it lends naturally to comparing two of the approaches mentioned in [35]. That is, it lets us work with quantum states by inputting them as the physical state of the cells of the quantum tape, or alternatively, we can input their classical representation by encoding the corresponding density matrix as a string in the classical tape. We defined the algorithmic complexity of a state ρ as the size of the description of the shortest computable sequence of "outputtable" states that approximates ρ, and showed that it is machine independent. We also showed that any program that computes ρ can be transformed into one that computes $\boldsymbol{\rho}$, and vice-versa; which means that the resulting complexities $\mathrm{K}(\rho)$ and $\mathrm{K}(\boldsymbol{\rho})$ are equivalent. Nevertheless, we found that the conditional complexities $\mathrm{K}(\sigma|\rho)$ and $\mathrm{K}(\sigma|\boldsymbol{\rho})$

[4] Models of computation that allow for performing fully quantum programs have the problem of outputting uncomputable quantum states. For instance, consider an uncomputable real number $p \in (0,1)$. According to the Physical Church-Turing postulate, there is no physical system able to sample the Bernoulli distribution with parameter p. Otherwise, we could approximate arbitrarily p, by the law of large numbers and sampling. However, fully quantum programs would be able to create this distribution by preparing the state $\sqrt{1-p}|0\rangle + \sqrt{p}|1\rangle$ and measuring it on the computational basis.

are, in general, not the same. Instead, the classical representations contain more descriptive information than the physical system when used as a resource. This led us to study the behaviour of complexity for quantum state copying. By combining Equations (31) and (39) we can obtain the relation

$$\mathrm{K}(\rho) = \mathrm{K}(\rho \otimes \rho \mid \rho) + O(n), \tag{94}$$

which can be seen as a statement of how much of an advantage having a copy of a state gives when computing two copies. For any qubit number n and value k of complexity, there is a finite number of quantum states for which $\mathrm{K}(\rho) \leq k$, and an infinite number of states for which $\mathrm{K}(\rho) > k$. It follows that for any $n \in \mathbb{N}$ and integer function $k(n) \in O(n)$, there is an infinite number of n-qubit states ρ for which $\mathrm{K}(\rho) \gg k(n)$. Therefore, for *almost all* quantum states, having access to a copy of the state gives no significant advantage for computing its duplicate.

Finally, we considered notions of quantum algorithmic mutual information, and found that two classically equivalent definitions are no longer equivalent for the quantum case. We noted that the property that connects these two quantities in the classical case, namely the chain rule, does not hold for general bipartite quantum states. We interpreted this discrepancy to be caused by the presence of quantum correlations. To compute a quantum multipartite state it is not enough to be able to compute each of its parts individually, but one must also describe the inherent correlations present in the joint state. Taking this into account, we developed a generalization of the chain rule for quantum states in the dcq-TM model. This new property let us to propose a candidate for a measure of the algorithmic complexity of correlations in quantum systems, which is symmetric up to a logarithmic term in the complexity of the joint state times the number of qubits.

Potential future work directions include the study of the relationship between quantum Kolmogorov complexity and Von Neumann entropy, one way to approach this is through generalization of Brudno's Theorem, which relates the complexity rate with the entropy rate for stationary, ergodic sources. Such analysis has already been done in [3] for Deutsch's machine model [8], which is quantum-controlled. Additionally, further properties of complexity of correlations can be explored and a potential theory of algorithmic complexity of multipartite correlations can be developed.

Acknowledgements

The research on this paper was funded under the FCT project QuantumPrime reference: PTDC/EEI-TEL/8017/2020 (50%) the QuantaGENOMICS project, through the EU H2020 QuantERA II Programme, Grant Agreement No 101017733, CERN/FIS-PAR/0023/2019 (50%). The authors acknowledge Fundação para a Ciência e Tecnologia, Instituto de Telecomunicações Research Unit, ref. UIDB/50008/2020, UIDP/50008/2020 and PEst-OE/EEI/LA0008/2013 and LASIGE Research Unit, ref. UIDB/00408/2020 and ref. UIDP/00408/2020. NP acknowledges the FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT000. ML acknowledges the PhD scholarship PD/BD/114334/2016.

References

[1] L. Antunes, A. Matos, A. Pinto, A. Souto, and A. Teixeira. One-way functions using algorithmic and classical information theories. *Theory of Computing Systems*, 52(1):162–178, Jan 2013. ISSN 1433-0490. DOI: 10.1007/s00224-012-9418-z.

[2] D. Azevedo, A. M. Rodrigues, H. Canhão, A. M. Carvalho, and A. Souto. Zgli: A pipeline for clustering by compression with application to patient stratification in spondyloarthritis. *Sensors*, 23(3), 2023. ISSN 1424-8220. DOI: 10.3390/s23031219.

[3] F. Benatti, T. Krüger, M. Müller, R. Siegmund-Schultze, and A. Szkoła. Entropy and quantum Kolmogorov complexity: A quantum Brudno's theorem. *Commun. Math. Phys.*, 265(1):437–461, 2006. DOI: 10.1007/s00220-006-0027-z.

[4] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, page 175, 1984. DOI: 10.1016/j.tcs.2014.05.025.

[5] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5): 1411–1473, 1997. DOI: 10.1137/S0097539796300921.

[6] A. Berthiaume, W. Dam, and S. Laplante. Quantum Kolmogorov complexity. *Journal of Computer and System Sciences*, 63(2):201–221, 2001. DOI: 10.1006/jcss.2001.1765.

[7] G. Chaitin. On the length of programs for computing finite binary sequences. *J. ACM*, 13(4), 1966. DOI: 10.1145/321356.321363.

[8] D. Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Royal Society of London Proceedings Series A*, 400(1818):97–117, 1985. DOI: 10.1098/rspa.1985.0070.

[9] P. Gács. Quantum algorithmic entropy. *Journal of Physics A: Mathematical and General*, 34(35): 6859, 2001. DOI: 10.1088/0305-4470/34/35/312.

[10] Peter Grünwald and Paul Vitányi. *Algorithmic Information Theory*, pages 289–325. E, January 2008. DOI: 11245/1.297540.

[11] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of modern physics*, 81(2):865, 2009. DOI: 10.1103/RevModPhys.81.865.

[12] A. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1(1), 1965. DOI: 10.1080/00207166808803030.

[13] T. Lee and A. Romashchenko. Resource bounded symmetry of information revisited. *Theoretical Computer Science*, 345(2):386–405, 2005. ISSN 0304-3975. DOI: 10.1016/j.tcs.2005.07.017. Mathematical Foundations of Computer Science 2004.

[14] Ming Li and Paul M. B. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications, 4th Edition*. Texts in Computer Science. Springer, 2019. ISBN 978-3-030-11297-4. DOI: 10.1007/978-3-030-11298-1.

[15] Noah Linden and Sandu Popescu. The halting problem for quantum computers. *arXiv preprint quant-ph/9806054*, 1998. DOI: 10.48550/arXiv.quant-ph/9806054.

[16] P. Mateus, A. Sernadas, and A. Souto. Universality of quantum Turing machines with deterministic control. *Journal of Logic and Computation*, 27(1):1–19, 2017. DOI: 10.1093/logcom/exv008.

[17] T. Miyadera. Quantum Kolmogorov complexity and information-disturbance theorem. *Entropy*, 13 (4):778–789, 2011. ISSN 1099-4300. DOI: 10.3390/e13040778.

[18] T. Miyadera and H. Imai. Quantum Kolmogorov complexity and quantum key distribution. *Phys. Rev. A*, 79:012324, Jan 2009. DOI: 10.1103/PhysRevA.79.012324.

[19] Takayuki Miyadera and Masanori Ohya. On halting process of quantum turing machine. *Open Systems & Information Dynamics*, 12(3):261–264, 2005. DOI: 10.1007/s11080-005-0923-2.

[20] Kavan Modi, Aharon Brodutch, Hugo Cable, Tomasz Paterek, and Vlatko Vedral. The classical-quantum boundary for correlations: Discord and related measures. *Reviews of Modern Physics*, 84 (4):1655, 2012. DOI: 10.1103/RevModPhys.84.1655.

[21] C. Mora and H. Briegel. Algorithmic complexity and entanglement of quantum states. *Physical Review Letters*, 95:200503, 2005. DOI: 10.1103/PhysRevLett.95.200503.

[22] C. Mora, H. Briegel, and B. Kraus. Quantum Kolmogorov complexity and its applications. *International Journal of Quantum Information*, 2007. DOI: 10.1142/S0219749907003171.

[23] M Muller. *Quantum Kolmogorov complexity and the quantum Turing machine*. Ph.D. Thesis, Technical University of Berlin, 2007. DOI: 10.48550/arXiv.0712.4377.

[24] M. Müller. Strongly universal quantum Turing machines and invariance of Kolmogorov complexity. *IEEE Transactions on Information Theory*, 54(2):763–780, 2008. ISSN 0018-9448. DOI: 10.1109/TIT.2007.913263.

[25] John M Myers. Can a universal quantum computer be fully quantum? *Physical Review Letters*, 78 (9):1823, 1997. DOI: 10.1103/PhysRevLett.78.1823.

[26] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511976667.

[27] A Rastegin. A lower bound on the relative error of mixed-state cloning and related operations. *Journal of Optics B: Quantum and Semiclassical Optics*, 5(6):S647, 2003. DOI: 10.1088/1464-4266/5/6/017.

[28] A. Sarkar, Z. Al-Ars, and K. Bertels. Estimating algorithmic information using quantum computing for genomics applications. *Applied Sciences*, 11(6), 2021. ISSN 2076-3417. DOI: 10.3390/app11062696.

[29] Claude Elwood Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 7 1948. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

[30] R. Solomonoff. A formal theory of inductive inference, part i. *Information and Control*, 7(1), 1964. DOI: 10.1016/S0019-9958(64)90223-2.
[31] A. Souto, L. Antunes, P. Mateus, and A. Teixeira. Witness hiding without extractors or simulators. In F. Manea, R. Miller, and D. Nowotka, editors, *Sailing Routes in the World of Computation*, pages 397–409, Cham, 2018. Springer International Publishing. DOI: 10.1007/978-3-319-94418-0˙40.
[32] K. Svozil. Quantum algorithmic information theory. *Journal of Universal Computer Science*, 2(5): 311–346, may 1996. DOI: 10.3217/jucs-002-05-0311.
[33] Andreia Teixeira, Armando Matos, André Souto, and Luís Antunes. Entropy measures vs. Kolmogorov complexity. *Entropy*, 13(3):595–611, 2011. ISSN 1099-4300. DOI: 10.3390/e13030595.
[34] P. Vitányi. Quantum Kolmogorov complexity based on classical descriptions. *IEEE Transactions on Information Theory*, 47(6):2464–2479, 2001. DOI: 10.1109/18.945258.
[35] Paul Vitanyi. Three approaches to the quantitative definition of information in an individual pure quantum state. In *Proceedings 15th Annual IEEE Conference on Computational Complexity*, pages 263–270. IEEE, 2000. DOI: 10.1109/CCC.2000.856757.
[36] A K Zvonkin and L A Levin. The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematical Surveys*, 25(6):83, dec 1970. DOI: 10.1070/RM1970v025n06ABEH001269.

8 Appendix

8.1 Proof of Lemma 5.4

Definition 8.1. *The Bures angle $\Delta(\rho_1, \rho_2)$ between two quantum states ρ_1, ρ_2 is defined as:*

$$\Delta(\rho_1, \rho_2) = \cos^{-1}(\sqrt{F(\rho_1, \rho_2)}) \tag{95}$$

The Bures angle is a metric in the space of density operators acting on a given Hilbert space. Moreover, it is contractive under quantum operations [26], that is, for any CPTP transformation T it holds that

$$\Delta(T(\rho_1), T(\rho_2)) \leq \Delta(\rho_1, \rho_2). \tag{96}$$

Lemma 8.1. *Let $T : \mathcal{D}(\mathcal{H}) \to \mathcal{D}(\mathcal{H}^{\otimes 2})$ be a CPTP map. Define the* copying error $\texttt{CE}_T$ *of the transformation T on the state $\rho \in \mathcal{D}(\mathcal{H})$ as*

$$\texttt{CE}_T(\rho) = \sqrt{1 - F(\rho^{\otimes 2}, T(\rho))}. \tag{97}$$

For any $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ such that $\texttt{CE}_T(\rho_1) + \texttt{CE}_T(\rho_2) \leq \varepsilon \leq \frac{1}{4}$ it holds that

$$F(\rho_1, \rho_2) \leq \frac{1}{2} - \sqrt{\frac{1}{4} - \varepsilon} \quad \text{or} \quad F(\rho_1, \rho_2) \geq \frac{1}{2} + \sqrt{\frac{1}{4} - \varepsilon}. \tag{98}$$

Proof. From the triangle inequality of the Bures angle we know that

$$\Delta(\rho_1^{\otimes 2}, \rho_2^{\otimes 2}) \leq \Delta(\rho_1^{\otimes 2}, T(\rho_1)) + \Delta(T(\rho_1), T(\rho_2)) + \Delta(T(\rho_2), \rho_2^{\otimes 2}). \tag{99}$$

Rearranging terms and taking the sin on both sides we get

$$\sin\left(\Delta(\rho_1^{\otimes 2}, T(\rho_1)) + \Delta(T(\rho_2), \rho_2^{\otimes 2})\right) \geq \sin\left(\Delta(\rho_1^{\otimes 2}, \rho_2^{\otimes 2}) - \Delta(T(\rho_1), T(\rho_2))\right). \tag{100}$$

We can bound the left hand side of Equation (100) by recalling that

$$\sin\left(\Delta(\rho_1^{\otimes 2}, T(\rho_1)) + \Delta(T(\rho_2), \rho_2^{\otimes 2})\right) \leq \underbrace{\sin\left(\Delta(\rho_1^{\otimes 2}, T(\rho_1))\right)}_{\texttt{CE}_T(\rho_1)} + \underbrace{\sin\left(\Delta(T(\rho_2), \rho_2^{\otimes 2})\right)}_{\texttt{CE}_T(\rho_2)}. \tag{101}$$

Additionally, noting that

$$\begin{aligned}\sin\left(\Delta(\rho_1^{\otimes 2}, \rho_2^{\otimes 2}) - \Delta(T(\rho_1), T(\rho_2))\right) = &\sin\left(\Delta(\rho_1^{\otimes 2}, \rho_2^{\otimes 2})\right)\cos\left(\Delta(T(\rho_1), T(\rho_2))\right) \\ &- \cos\left(\Delta(\rho_1^{\otimes 2}, \rho_2^{\otimes 2})\right)\sin\left(\Delta(T(\rho_2), \rho_2^{\otimes 2})\right),\end{aligned} \tag{102}$$

and using the following identities/inequalities (derived from the contractive property of the Bures angle and the squaring property of fidelity) we can bound the right hand side of Equation (100) and write it in terms of fidelity:

$$\begin{aligned}
\sin\left(\Delta(\rho_1^{\otimes 2}, \rho_2^{\otimes 2})\right) &= \sqrt{1 - F(\rho_1^{\otimes 2}, \rho_2^{\otimes 2})} = \sqrt{1 - F(\rho_1, \rho_2)^2} \\
\cos\left(\Delta(T(\rho_1), T(\rho_2))\right) &\geq \cos\left(\Delta(\rho_1, \rho_2)\right) = \sqrt{F(\rho_1, \rho_2)} \\
\cos\left(\Delta(\rho_1^{\otimes 2}, \rho_2^{\otimes 2})\right) &= \sqrt{F(\rho_1^{\otimes 2}, \rho_2^{\otimes 2})} = F(\rho_1, \rho_2) \\
\sin\left(\Delta(T(\rho_1), T(\rho_2))\right) &\leq \sin\left(\Delta(\rho_1, \rho_2)\right) = \sqrt{1 - F(\rho_1, \rho_2)}.
\end{aligned} \tag{103}$$

Substituting the inequalities (101) and (103) in the left and right hand sides of Equation (100), respectively, we get:

$$\begin{aligned}
\mathtt{CE}_T(\rho_1) + \mathtt{CE}_T(\rho_2) &\geq \sqrt{(F(\rho_1, \rho_2))(1 - F(\rho_1, \rho_2)^2)} + F(\rho_1, \rho_2)\sqrt{1 - F(\rho_1, \rho_2)} \\
&\geq F(\rho_1, \rho_2)(1 - F(\rho_1, \rho_2)),
\end{aligned} \tag{104}$$

where the last step holds because $F(\rho_1, \rho_2) \in [0, 1]$. Finally, assume that $\mathtt{CE}_T(\rho_1) + \mathtt{CE}_T(\rho_2) \leq \varepsilon \leq \frac{1}{4}$, then the solutions of the quadratic inequality in the interval $[0, 1]$ satisfy

$$\varepsilon \geq F(\rho_1, \rho_2)(1 - F(\rho_1, \rho_2)) \implies F(\rho_1, \rho_2) \leq \frac{1}{2} - \sqrt{\frac{1}{4} - \varepsilon} \quad \text{or} \quad F(\rho_1, \rho_2) \geq \frac{1}{2} + \sqrt{\frac{1}{4} - \varepsilon}. \tag{105}$$

□

8.2 Proof of Lemma 5.5

Lemma 5.5 is a generalization of the basic property which states that for a finite dimensional Hilbert space $\mathcal{H}$, any set $\{\rho_i\}$ of orthogonal states can have at most $\dim(\mathcal{H})$ elements. The lemma states that this property holds even if the states are not quite orthogonal, but instead have pair-wise fidelities upper bounded by $\delta_0(\mathcal{H})$. To prove this result we will start by proving a restricted version of the main lemma, which is then used to prove the more general statement. First, we restrict the sets to consist only of pure states, then we build upon that version removing the requirement of the states to be pure.

Theorem 8.2. *(Gram-Schmidt orthogonalization) Let $\mathcal{H}$ be a Hilbert space and $\{|\psi_i\rangle \in \mathcal{H}\}_{i=1}^k$ a normalized, linearly independent set. For $i = 1, 2, \ldots, k$ define*

$$|\tilde{e}_i\rangle = |\psi_i\rangle - \sum_{j<i} \langle \tilde{e}_j | \psi_i \rangle |\tilde{e}_j\rangle, \tag{106}$$

and

$$|e_i\rangle = \frac{|\tilde{e}_i\rangle}{||\tilde{e}_i\rangle|}. \tag{107}$$

The set $\{|e_i\rangle\}_{i=1}^k$ is an orthonormal set.

The following properties refer directly to the sets defined in Theorem 8.2:

Lemma 8.3. *Let $\mathcal{H}$ be an N-dimensional Hilbert space, $0 \leq \delta \leq 1$, and $\{|\psi_i\rangle \in \mathcal{H}\}$ a normalized, linearly independent set such that for $i \neq j$*

$$|\langle \psi_i | \psi_j \rangle|^2 \leq \delta. \tag{108}$$

The non-normalized Gram-Schmidt vectors $\{|\tilde{e}_i\rangle\}_{i=1}^k$ satisfy for $j < i$

$$|\langle \psi_i | \tilde{e}_j \rangle| \leq a_{j-1}\sqrt{\delta}, \tag{109}$$

where

$$a_j = \begin{cases} 1 & \text{if } j = 0 \\ a_{j-1}^2 + a_{j-1} & \text{if } j \geq 1 \end{cases} \tag{110}$$

Proof. We proceed by induction over the index j. Note that for $j = 1$ and any $i > 1$ we have that

$$|\langle\psi_i|\tilde{e}_j\rangle| = |\langle\psi_i|\psi_1\rangle| \leq \sqrt{\delta} = a_0\sqrt{\delta}. \tag{111}$$

Assume now that for some i and $j < i - 1$ it holds that for all $1 < i' \leq i, 1 \leq j' \leq j$ and $j' < i'$

$$|\langle\psi_{i'}|\tilde{e}_{j'}\rangle| \leq a_{j'-1}\sqrt{\delta}, \tag{112}$$

then

$$\begin{aligned}
|\langle\psi_i|\tilde{e}_{j+1}\rangle| &= \left|\langle\psi_i|\psi_{j+1}\rangle - \sum_{k\leq j}\langle\tilde{e}_k|\psi_{j+1}\rangle\langle\psi_i|\tilde{e}_k\rangle\right| \\
&\leq \underbrace{\left|\langle\psi_i|\psi_{j+1}\rangle\right|}_{\leq\sqrt{\delta}} + \sum_{k=1}^{j}\underbrace{\left|\langle\tilde{e}_k|\psi_{j+1}\rangle\right|\left|\langle\psi_i|\tilde{e}_k\rangle\right|}_{\leq a_{k-1}^2\delta \;\leq\; a_{k-1}^2\sqrt{\delta}} \\
&\leq \left(1 + \sum_{k=1}^{j} a_{k-1}^2\right)\sqrt{\delta} \\
&= a_j\sqrt{\delta}
\end{aligned} \tag{113}$$

□

Corolary 8.4. *Under the assumptions of Lemma 8.3, with* $\delta \leq 2^{-3^N}$

$$|\langle\psi_i|e_i\rangle| \geq \left|1 - 2^{3^N}\delta\right| \tag{114}$$

Proof. From Equation (106) we know that

$$|\psi_i\rangle = |\tilde{e}_i\rangle + \sum_{j<i}\langle\tilde{e}_j|\psi_i\rangle|\tilde{e}_j\rangle. \tag{115}$$

Recalling that the two terms on the right side are mutually orthogonal and that $||\psi_i\rangle| = 1$, it follows that $||\tilde{e}_i\rangle| \leq 1$ and therefore

$$|\langle\psi_i|e_j\rangle| \geq |\langle\psi_i|\tilde{e}_j\rangle|. \tag{116}$$

On the other hand, we can bound the sequence defined in Equation (114) by the doubly exponential sequence

$$a_{j+1} = a_j^2 + a_j \leq 2a_j^2 \leq 2^{3^j}, \tag{117}$$

and hence, by Lemma 8.3

$$\begin{aligned}
|\langle\psi_i|e_i\rangle| \geq |\langle\psi_i|\tilde{e}_i\rangle| &\geq \left|1 - \sum_{j<i}|\langle\tilde{e}_j|\psi_i\rangle|^2\right| \\
&\geq \left|1 - (a_i - 1)\delta\right| \\
&\geq \left|1 - (a_n)\delta\right| \\
&\geq 1 - 2^{3^N}\delta
\end{aligned} \tag{118}$$

□

Lemma 8.5. *(Pure states version) Let* $\mathcal{H}$ *be an* N*-dimensional Hilbert space and* $A = \{|\psi_i\rangle \in \mathcal{H}\}$ *a normalized set such that for* $i \neq j$

$$|\langle\psi_i|\psi_j\rangle|^2 \leq \delta_0^{\text{pure}}(N). \tag{119}$$

with

$$\delta_0^{\text{pure}}(N) = \left(\frac{1 - \sqrt{1 - \frac{1}{N}}}{1 + 2^{\left(\frac{3^N+1}{2}\right)}}\right)^2. \tag{120}$$

Then A *cannot have more than* N *elements.*

Proof. We intend to show that for $\delta_0^{\text{pure}}(N)$ defined as in Equation (120), the set A must be linearly independent. Let $A' = \{|\phi_i\rangle\}_{i=1}^M \subseteq A$ be a maximal linearly independent subset of A and $A'' = A \setminus A'$. We proceed to prove that any other state belonging to the Hilbert space $\mathcal{H}' = \text{Span}(A) = \text{Span}(A')$ cannot satisfy Equation (119), and thus A'' must be the empty set.

Let $\{|e_i\rangle\}_{i=1}^M$ be the set obtained from A' through the Gram-Schmidt process. Because the $|e_i\rangle$ form an orthonormal basis for $\mathcal{H}'$, it holds that for any $|\tilde{\phi}\rangle \in A'' \subseteq \mathcal{H}'$ there exists k such that

$$|\langle\tilde{\phi}|e_k\rangle|^2 \geq \frac{1}{M} \geq \frac{1}{N}. \tag{121}$$

Furthermore, from the triangle inequality of the trace distance

$$D(|\tilde{\phi}\rangle, |e_k\rangle) + D(|e_k\rangle, |\phi_k\rangle) \geq D(|\tilde{\phi}\rangle, |\phi_k\rangle), \tag{122}$$

writing the trace distances in terms of inner products we obtain

$$\sqrt{1 - \left|\langle\tilde{\phi}|e_k\rangle\right|^2} + \sqrt{1 - \left|\langle e_k|\phi_k\rangle\right|^2} \geq \sqrt{1 - \left|\langle\tilde{\phi}|\phi_k\rangle\right|^2}. \tag{123}$$

We can now use the inequalities (114) and (121) to get

$$\sqrt{1 - \frac{1}{N}} + \sqrt{1 - \left(1 - 2^{3^N}\delta_0^{\text{pure}}\right)^2} \geq \sqrt{1 - \left|\langle\tilde{\phi}|\phi_k\rangle\right|^2}, \tag{124}$$

which we can solve for $\left|\langle\tilde{\phi}|\phi_k\rangle\right|^2$ as follows

$$\left|\langle\tilde{\phi}|\phi_k\rangle\right|^2 \geq 1 - \left(\sqrt{1 - \frac{1}{N}} + \sqrt{1 - \left(1 - 2^{3^N}\delta_0^{\text{pure}}\right)^2}\right)^2. \tag{125}$$

All that remains now is to show that the right side of Equation (125) is strictly greater than δ_0^{pure}. Indeed, setting $a = 1 - \sqrt{1 - \frac{1}{N}}$ and $b = 2^{3^N}$ we can rewrite Equation (120) as

$$\sqrt{\delta_0^{\text{pure}}} = a - \sqrt{2b\delta_0^{\text{pure}}}, \tag{126}$$

and we can now use the fact that for all $N \geq 1$ it holds that $0 < \delta_0^{\text{pure}}(N) \leq 2^{-3^N} < 1$ to conclude that

$$\begin{aligned}
\delta_0^{\text{pure}} < \sqrt{\delta_0^{\text{pure}}} &= a - \sqrt{2b\delta_0^{\text{pure}}} \\
&< a - \sqrt{b\delta_0^{\text{pure}}(2 - b\delta_0^{\text{pure}})} \\
&= 1 - \underbrace{\left((1-a) + \sqrt{b\delta_0^{\text{pure}}(2 - b\delta_0^{\text{pure}})}\right)}_{\leq 1} \\
&\leq 1 - \left((1-a) + \sqrt{b\delta_0^{\text{pure}}(2 - b\delta_0^{\text{pure}})}\right)^2 \\
&= 1 - \left(\sqrt{1 - \frac{1}{N}} + \sqrt{1 - \left(1 - 2^{3^N}\delta_0^{\text{pure}}\right)^2}\right)^2 \\
&\leq \left|\langle\tilde{\phi}|\phi_k\rangle\right|^2.
\end{aligned} \tag{127}$$

This means that no state $|\tilde{\phi}\rangle \in \mathcal{H}'$ can satisfy both $|\tilde{\phi}\rangle \notin A'$ and $|\tilde{\phi}\rangle \in A$. Therefore, $A = A'$ is a linearly independent set and as such may have at most N elements. □

Lemma 8.6. *(General statement) Let $\mathcal{H}$ be an N-dimensional Hilbert space and $A = \{\rho_i \in \mathcal{D}(\mathcal{H})\}$ be a set of density operators such that for $i \neq j$:*

$$F(\rho_i, \rho_j) \leq \delta_0(N), \tag{128}$$

with

$$\delta_0(N) = \left(\frac{\delta_0^{\text{pure}}(N)}{N}\right)^2. \tag{129}$$

Then A cannot have more than N elements.

Proof. To prove the result we will proceed as follows: for each $\rho_i \in A$ we will define an associated pure state such that if A satisfies Equation (128) then the associated set I of pure states satisfies Equation (119), which in turn by lemma 8.5 implies that $|I| = |A| \leq N$.

Let $M = |A|$. Following the spectral decomposition theorem; for $i = 1, \ldots, M$, let $\{|\psi_i^k\rangle\}_{k=1}^N$ be a set of orthogonal states and $\{\lambda_i^k\}_{k=1}^N$ a set of non-negative real numbers such that

$$\rho_i = \sum_{k=1}^{N} \lambda_i^k |\psi_i^k\rangle\langle\psi_i^k|. \tag{130}$$

Furthermore, it is easy to check that for $i, j = 1, \ldots, M$ the following is an orthonormal set:

$$\left\{|\psi_j^s(i)\rangle = \sum_{k=1}^{N} \langle\psi_j^s|\psi_i^k\rangle|\psi_i^k\rangle\right\}_{s=1}^{N}. \tag{131}$$

We intend now to use the Uhlmann's theorem to find a bound for inner products $\langle\psi_j^s|\psi_i^k\rangle$ in terms of $\delta_0(N)$. Recall that any state of the form

$$\sum_{k=1}^{N} \sqrt{\lambda_i^k}|\psi_i^k\rangle_{\mathcal{H}}|\phi_i^k\rangle_{\mathcal{H}'} \tag{132}$$

is a purification of ρ_i in $\mathcal{H} \otimes \mathcal{H}'$ for any orthogonal basis $\{|\phi_i^k\rangle \in \mathcal{H}'\}$, where $\mathcal{H}'$ is an ancilliary Hilbert space isomorphic to $\mathcal{H}$. Now, from Ulhmann's theorem:

$$\begin{aligned}
\delta_0(N) \geq F(\rho_i, \rho_j) &= \max_{\Psi_i, \Psi_j} |\langle\Psi_i|\Psi_j\rangle|^2 \\
&\geq \Big| \sum_k \sqrt{\lambda_i^k}_{\mathcal{H}}\langle\psi_i^k|_{\mathcal{H}'}\langle\psi_i^k| \sum_s \sqrt{\lambda_j^s}|\psi_j^s\rangle_{\mathcal{H}}|\psi_j^s(i)\rangle_{\mathcal{H}'} \Big|^2 \\
&= \Big| \sum_{k,s} \sqrt{\lambda_i^k \lambda_j^s}\langle\psi_i^k|\psi_j^s\rangle \underbrace{\langle\psi_i^k|\psi_j^s(i)\rangle}_{\langle\psi_j^s|\psi_i^k\rangle} \Big|^2 \\
&= \Big| \sum_{k,s} \sqrt{\lambda_i^k \lambda_j^s}|\langle\psi_i^k|\psi_j^s\rangle|^2 \Big|^2,
\end{aligned} \tag{133}$$

where the maximum is taken over all purifications Ψ_i, Ψ_j of ρ_i, ρ_j, respectively. Since all terms in the last summation are positive numbers we can conclude that for any $k, s = 1, \ldots, N$:

$$\sqrt{\delta_0(N)} \geq \sqrt{\lambda_i^k \lambda_j^s}\big|\langle\psi_i^k|\psi_j^s\rangle\big|^2. \tag{134}$$

Define now the sets

$$I_i = \left\{|\psi_i^k\rangle \mid \lambda_i^k \geq \frac{1}{N}\right\}; \quad I = \bigcup_{i=1}^{M} I_i, \tag{135}$$

notice that the set I must have *at least* M elements because every density operator acting on an N-dimensional space must have at least one eigenvalue greater or equal to $\frac{1}{N}$. Substituting $|\psi_i^k\rangle \in I_i$ and $|\psi_j^s\rangle \in I_j$ for $i \neq j$ in Equation (134) we obtain

$$\frac{\delta_0^{\text{pure}}(N)}{N} = \sqrt{\delta_0(N)} \geq \sqrt{\lambda_i^k \lambda_j^s}\big|\langle\psi_i^k|\psi_j^s\rangle\big|^2 \geq \frac{1}{N}\big|\langle\psi_i^k|\psi_j^s\rangle\big|^2, \tag{136}$$

and thus

$$\delta_0^{\text{pure}}(N) \geq \big|\langle\psi_i^k|\psi_j^s\rangle\big|^2. \tag{137}$$

This means that the set I satisfies the conditions of lemma 8.5 and hence may have *at most* N elements. We can conclude then

$$M \leq |I| \leq N. \tag{138}$$

□

8.3 Enumeration of the Σ_s

In the proof of Theorem 6.2 the sets

$$\Sigma_s = \{t \in P_s |\ |{}_2P_s^{\eta(t,s)}| > 2^\ell\}, \tag{139}$$

were defined and claimed to be computably enumerable. Algorithm 5 provides a way to enumerate them by calling a computable enumeration $\mathcal{E}_s$ of the set P_s, and subroutine Υ, which on input (s, t, i):

- Outputs the i-th element of some fixed computable enumeration $\Upsilon_{s,t}$ of the set ${}_2P_s^{\eta(t,s)}$ if $i \leq |{}_2P_s^{\eta(t,s)}|$.
- Does not halt if $i > |{}_2P_s^{\eta(t,s)}|$.

Note that a subroutine with the above properties can always be constructed from a program that enumerates a set, since such program can be simulated to produce a certificate of correctness for its output whenever the index is less or equal to the size of the set (by printing the transcript of all its computation), so whenever it halts with an invalid certificate, the simulator can be made to enter an infinite cycle.

Parameters: Subroutine Υ, which in input (s, t, i) outputs the i-th element of some fixed computable enumeration $\Upsilon_{s,t}$ of the set ${}_2P_s^{\eta(t,s)}$. Computable enumeration $\mathcal{E}_s$ of the set P_s. Positive integers s, ℓ ;
Input: Integer $i \geq 1$;
$L \leftarrow (), j \leftarrow 1$;
while $|L| < i$ **do**
  Set n equal to the smallest natural number such that $j < \frac{(n+1)(n+2)}{2}$;
  if $n = 1$ **then**
    $m = s - n$;
  else
    $m = s \mod \frac{n(n+1)}{2}$;
  end
  $t' \leftarrow \mathcal{E}_s(n - m)$;
  Simulate the first m steps of $\Upsilon(s, t', 2^l + 1)$;
  if *The simulated machine reached its halting state and* t' *is not in* L **then**
    Add t' to L
  end
  $j \leftarrow j + 1$;
end
return $L[i]$

Algorithm 5: Enumeration of the set Σ_s given computable enumerations $\mathcal{E}_s, \Upsilon_{s,t}$ of the respective sets P_s and ${}_2P_s^{\eta(t,s)}$, and the values of s and ℓ.

Operational interpretation of the vacuum and process matrices for identical particles

Ricardo Faleiro[1], Nikola Paunković[1,2], and Marko Vojinović[3]

[1]Instituto de Telecomunicações, Avenida Rovisco Pais 1, 1049-001, Lisboa, Portugal
[2]Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001, Lisboa, Portugal
[3]Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

arXiv:2010.16042v5 [quant-ph] 14 Apr 2023

This work overviews the single-particle two-way communication protocol recently introduced by del Santo and Dakić (dSD), and analyses it using the process matrix formalism. We give a detailed account of the importance and the operational meaning of the interaction of an agent with the vacuum — in particular its role in the process matrix description. Our analysis shows that the interaction with the vacuum should be treated as an operation, on equal footing with all other interactions. This raises the issue of counting such operations in an operational manner. Motivated by this analysis, we apply the process matrix formalism to capped Fock spaces using the framework of second quantisation, in order to characterise protocols with an indefinite number of identical particles.

1 Introduction

In recent years there have been advances in quantum information theory related to new techniques for discussing quantum circuits and quantum computation. One of those techniques is the recently developed process matrix formalism [1]. This formalism is general enough to describe all known quantum processes, in particular the superposed orders of operations in a quantum circuit. Moreover, its most prominent feature is that it allows for a description of more general situations of indefinite causal orders of spacetime points. A formal example of such a process has been introduced and discussed in [1], leading to the violation of the so-called *causal inequalities*. The latter represent device-independent conditions that need to be satisfied in order for a given process to have a well-defined causal order. It is still an open question whether such a process is physical and can be realised in nature. Also, a lot of attention in the literature has been devoted to the *quantum switch* operation, which has been discussed through both theoretical descriptions [2, 3, 4, 5] and experimental implementations [6, 7, 8].

One of the interesting aspects of the quantum switch is that it gives rise to a superposition of orders of quantum operations. In a recent work [9], the difference between the superposition of orders of quantum operations and the superposition of causal orders in spacetime was discussed in detail, and it was demonstrated that the latter can in principle be realised only in the context of quantum gravity (see also [10, 11, 12]). The detailed analysis of the causal structure of the quantum switch has revealed one important qualitative aspect of the process matrix description — in order to properly account for the causal structure of an arbitrary process, it is *necessary* to introduce the notion of the quantum vacuum as a possible physical state. Otherwise, the naive application of the process matrix formalism may suggest a misleading conclusion that quantum switch implementations in flat spacetime feature genuine superpositions of spacetime causal orders. This demonstrates the importance of the concept of vacuum in quantum information processing. Regarding the general role of the vacuum in quantum circuits and optical experiments, see [13] and [14, 15], respectively, and the references therein.

Ricardo Faleiro: ricardofaleiro@tecnico.ulisboa.pt
Nikola Paunković: npaunkov@math.tecnico.ulisboa.pt
Marko Vojinović: vmarko@ipb.ac.rs

Simultaneously with these developments, another interesting quantum process has been recently proposed [16] by del Santo and Dakić — dSD protocol (see also subsequent theoretical [17, 18] and experimental work [19]). As it turns out, while this process enables Alice and Bob to guess each other's input bits with certainty by exchanging a single particle only once, it cannot be correctly described within the process matrix formalism without the introduction of the interaction between the vacuum and the apparatus as an *operation*. Thus, it represents an additional motivation to introduce the vacuum state into the process matrix formalism, independent of the reasons related to the description of the quantum switch process.

Moreover, while the dSD protocol employs only one photon, it is also relevant for multiphoton processes, which opens the question of the treatment of identical particles within the process matrix formalism. Also, taking into account the presence of the vacuum state, one is steered towards the application of the abstract process matrix formalism to systems with variable number of identical particles, to the second quantisation and ultimately generalisation to quantum field theory (QFT). See also a related work on the causal boxes framework [20].

In this work we give a detailed description and treatment of dSD protocol within the process matrix formalism. We analyse in detail the role of the vacuum in the protocol and the formalism, and its operational interpretation. Specifically, our aim is to discuss the following question:

Is the interaction with the vacuum an operation, or not?

Our analysis of dSD protocol leads to a conclusion that the interaction with the vacuum should be considered an operation. The alternative would mean that one could extract information from the system at the final moment of the protocol without performing an operation at all. Since the same physical situation should always be described in the same way, we conclude that the interaction with the vacuum should be treated as an operation, and thus as a resource, in all quantum information protocols. This includes the optical implementation of the quantum switch protocol, leading one to infer that it features *four, rather than just two* operations, as was claimed in a number of papers [3, 4, 5, 6, 7, 8]. In addition, we make use of the dSD protocol as an illuminative example to apply the process matrix formalism to multipartite systems of identical particles.

The paper is organised as follows. In Section 2 we give a short overview of the process matrix formalism and the dSD protocol. Section 3 is devoted to the process matrix formalism description of dSD protocol, and to the discussion of the *operational* role and importance of the vacuum state for its description. In particular, in Subsection 3.4 we present the argument for our main conclusion, namely that the interaction with the vacuum should be considered an operation. In Section 4 we provide the basic rules for the application of the process matrix formalism to identical multiparticle systems. Section 5 is devoted to the summary of our results, discussion and prospects for future research. The Appendix contains various technical details of the calculations.

2 State of the art

In this section, we present an overview of the relevant background results. First, we give a short introduction to the process matrix formalism, and then we present the dSD protocol. This overview is not intended to be complete or self-contained, but merely of informative type. The reader should consult the literature for more details.

2.1 The process matrix formalism

The process matrix formalism is based on an idea of a set of laboratories, interacting with the outside world by exchanging quantum systems. Each laboratory is assumed to be spatially local in the sense that one can consider its size negligible for the problem under discussion. Inside the laboratory, it is assumed that the ordinary laws of quantum theory hold. The laboratory interacts with the outside world by receiving an *input* quantum system and by sending an *output* quantum system. Inside the laboratory, the input and output quantum systems are being manipulated using the notion of an *instrument*, denoting the most general operation one can perform over quantum systems. Each interaction is also assumed to be localised in time, such that each operation of a given laboratory has a separate spacetime point

assigned to it (see Subsection 3.4 for a discussion of time delocalised laboratories and operations [21]). Thus, we introduce the notion of a *gate*, which represents the action of an instrument at a given spacetime point (see Section 2 of [9]); for simplicity, both the gate and its corresponding spacetime point will be denoted by the same symbol, G. By G_I and G_O we denote the Hilbert spaces of the input and the output quantum systems, respectively. These Hilbert spaces are assumed to be finite-dimensional or trivial. The action of the instrument is described by an operator, $\mathcal{M}^G_{x,a} : G_I \otimes G_I^* \to G_O \otimes G_O^*$, which may depend on some classical input information a available to the gate G, and some readouts x of eventual measurement results that may take place in G. Thus, the instrument maps a generic mixed input state ρ_I into the output state $\rho_O = \mathcal{M}^G_{x,a}(\rho_I)$.

Given such a setup, one defines a *process*, denoted $\mathcal{W}$, as a functional over the instruments of all gates, as

$$p(x, y, \dots | a, b, \dots) = \mathcal{W}(\mathcal{M}^{G^{(1)}}_{x,a} \otimes \mathcal{M}^{G^{(2)}}_{y,b} \otimes \dots),$$

where $p(x, y, \dots | a, b, \dots)$ represents the probability of obtaining measurement results $x, y, \dots$, given the inputs $a, b, \dots$. In order for the right-hand side to be interpreted as a probability distribution, the process $\mathcal{W}$ must satisfy three basic axioms,

$$\begin{aligned} \mathcal{W} &\geq 0\,, \\ \operatorname{Tr} \mathcal{W} &= \textstyle\prod_i \dim G_O^{(i)}\,, \\ \mathcal{W} &= \mathcal{P}_G(\mathcal{W})\,, \end{aligned} \tag{1}$$

where $\mathcal{P}_G$ is a certain projector onto a subspace of $\bigotimes_i \left(G_I^{(i)} \otimes G_O^{(i)}\right)$ which, together with the second requirement, ensures the normalisation of the probability distribution (see [3] for details).

In order to have a computationally manageable formalism, one often employs the Choi-Jamiołkowski (CJ) map over the instrument operations, such that a given operation $\mathcal{M}^G_{x,a}$ is being represented by a matrix,

$$\begin{aligned} M^G_{x,a} &= \left[\left(\mathcal{I} \otimes \mathcal{M}^G_{x,a}\right) (|\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}|) \right]^T \\ &\in (G_I \otimes G_O) \otimes (G_I \otimes G_O)^*, \end{aligned} \tag{2}$$

where

$$|\mathbb{1}\rangle\!\rangle \equiv \sum_i |i\rangle \otimes |i\rangle \in G_I \otimes G_I \tag{3}$$

is the so-called *transport vector*, representing the non-normalised maximally entangled state, and $\mathcal{I}$ is the identity operator. Then, one can describe the process $\mathcal{W}$ using the *process matrix* W to write

$$\begin{aligned} &p(x, y, \dots | a, b, \dots) = \\ &\qquad \operatorname{Tr}\left[(M^{G_1}_{x,a} \otimes M^{G_2}_{y,b} \otimes \dots) W \right]. \end{aligned} \tag{4}$$

Finally, if an instrument $\mathcal{M}^G_{x,a}$ is linear, one can also use a corresponding "vector" notation (see Appendix A.1 in [3]),

$$|(\mathcal{M}^G_{x,a})^*\rangle\!\rangle \equiv \left[\mathcal{I} \otimes (\mathcal{M}^G_{x,a})^*\right] |\mathbb{1}\rangle\!\rangle \in G_I \otimes G_O\,, \tag{5}$$

so that

$$M^G_{x,a} = |(\mathcal{M}^G_{x,a})^*\rangle\!\rangle\langle\!\langle (\mathcal{M}^G_{x,a})^*|\,. \tag{6}$$

In cases where all instruments are linear, and in addition the process matrix W is a one-dimensional projector, one can introduce the the corresponding *process vector* $|W\rangle\!\rangle$, such that $W = |W\rangle\!\rangle\langle\!\langle W|$, and rewrite (4) in the form:

$$\begin{aligned} &p(x, y, \dots | a, b, \dots) = \\ &\left\| \left(\langle\!\langle \mathcal{M}^{G^{(1)}}_{x,a} *| \otimes \langle\!\langle \mathcal{M}^{G^{(2)}}_{y,b} *| \otimes \dots \right) |W\rangle\!\rangle \right\|^2. \end{aligned} \tag{7}$$

2.2 The dSD protocol

In a recent paper [16], del Santo and Dakić have introduced a protocol which allows two agents to guess each other's input bits with certainty by exchanging a single particle only once. The protocol goes as follows. Initially, a single particle is prepared in a superposition state of being sent to Alice and being sent to Bob. Upon receiving the particle, both Alice and Bob perform unitary operations on it, encoding their bits of information, a and b, respectively, about the outcomes of their coin tosses. They do this by changing the local phase of the particle by $(-1)^a$ and $(-1)^b$. The particle is subsequently forwarded to a beam splitter, and after that again to Alice and Bob, who now measure the presence or absence of the particle.

This way, the state of the particle stays in *coherent* superposition of different paths in a Mach-Zehnder interferometer. The interference of its paths gives rise to *deterministic* outcome that depends on the relative phase $e^{i\phi} = (-1)^{a \oplus b}$ between the two branches: in case $\phi = 0$, the particle will end up in Alice's laboratory, while otherwise it will end in Bob's. Thus, knowing their own

inputs and the outcomes of their local measurements, both agents can determine each other's inputs, allowing for two-way communication using only one particle. This is clearly impossible in classical physics, demonstrating yet another example of the advantage of quantum over classical strategies.

The crucial aspect of the protocol lies in the fact that the *absence* of the particle represents a useful piece of information for an agent. This gives rise to the notion of the *vacuum state* as a carrier of information, playing the central role in the protocol. Thus, in order to describe the protocol using the process matrix formalism, one has to incorporate the notion of the vacuum in the formalism itself. We show this in detail in the next section.

It is interesting to note that the crucial role of the vacuum plays an important part not only in the dSD protocol, but also in a completely different setup that has been discussed a lot in recent literature, namely the *quantum switch* [2]. As analysed in detail in [9], if one takes care to distinguish the two temporal positions of a given laboratory and introduces the notion of a vacuum explicitly, one can demonstrate that the optical implementations of the quantum switch in flat spacetime do not feature any superposition of causal orders induced by the spacetime metric. Instead, it was argued that superpositions of spacetime causal orders can be present only within the context of a theory of quantum gravity. As we shall see below, the notion of the interaction with the vacuum will prove essential to the process matrix description of the dSD protocol as well.

3 Process matrix description of the dSD protocol

3.1 The spacetime diagram

We begin by drawing the spacetime diagram of the process corresponding to the dSD protocol (see Figure 1).

At the initial time t_i the laser L creates a photon and shoots it towards the beam splitter S, which at time t_1 performs the Hadamard operation and entangles it with the incoming vacuum state (described by the dotted arrow from the grey gate V). The entangled state of the photon and the vacuum continues on towards Alice's and Bob's gates A and B, respectively. At time t_2, Alice and Bob generate their random bits a and b, and encode them into the phase of the incoming photon-vacuum system. The system then proceeds to the beam splitter S' which again performs the Hadamard operation at time t_3. The photon-vacuum system then proceeds to the gates A' and B', where it is measured at time t_f by Alice and Bob, respectively. Note that the spatial distance Δl between Alice and Bob is precisely equal to the time distance between the generation of the random bits and the final measurements,

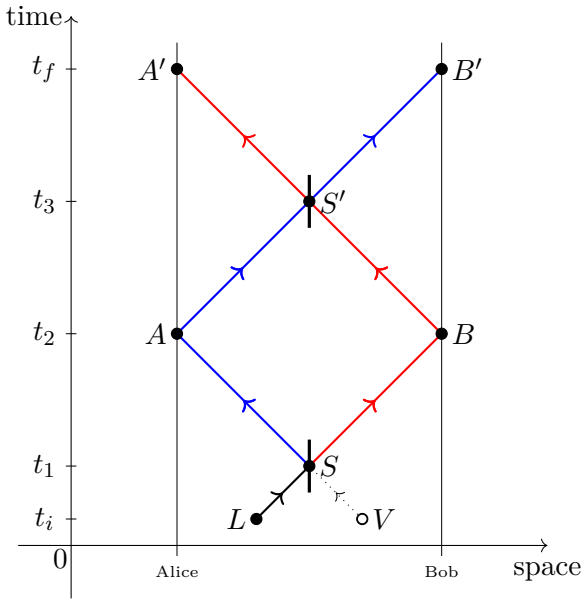


Figure 1: The complete spacetime diagram of the process corresponding to the dSD protocol.

$$\Delta l = c(t_f - t_2)\,,$$

so that a single photon has time to traverse the space between Alice and Bob *only once*. Also, note that the gate V, which generates the vacuum state, corresponds to a "trivial instrument", since the vacuum does not require any physical device to be generated. Nevertheless, the vacuum is still a legitimate physical state of the EM field, so the appropriate gate V has to be formally introduced and accounted for in the process matrix formalism calculations.

3.2 Formulation of the process matrix

Based on the spacetime diagram, we formulate the process matrix description as follows. All

spacetime points, where interaction between the EM field and some apparatus may happen, are assigned a gate and an operation which describes the interaction. Each gate has an input and output Hilbert space, as follows:

$$\begin{array}{lllll} L: & \mathbb{C} & \to & L_O\,, & \qquad A: \quad A_I \to A_O\,, \\ V: & \mathbb{C} & \to & V_O\,, & \qquad B: \quad B_I \to B_O\,, \end{array}$$

$$\begin{array}{lllll} S: & S_I & \to & S_O\,, & \qquad A': \quad A'_I \to \mathbb{C}\,, \\ S': & S'_I & \to & S'_O\,, & \qquad B': \quad B'_I \to \mathbb{C}\,. \end{array}$$

The initial gates, L and V, have trivial input spaces and nontrivial output spaces. The final gates, A' and B', have trivial output spaces and nontrivial input spaces. The gates A and B have nontrivial input and output spaces. Each nontrivial space is isomorphic to $\mathcal{H}_1 \oplus \mathcal{H}_0 \subset \mathcal{F}$, where $\mathcal{H}_0$ and $\mathcal{H}_1$ are the vacuum and single-excitation subspaces of the Fock space $\mathcal{F}$ in perturbative QED. Namely, by design of the dSD protocol, Alice and Bob may exchange at most one photon, which means that multiparticle Hilbert subspaces of the Fock space can be omitted. Moreover, the resulting probability distribution of the experiment outcomes does not in principle depend on the frequency or the polarisation of the photon in use, so we can approximate the single-excitation space as a one-dimensional Hilbert space. Given that the vacuum Hilbert space $\mathcal{H}_0$ is by definition one-dimensional, we can write

$$\mathcal{H}_0 = \mathrm{span}\{|0\rangle\} \equiv \mathbb{C}\,, \qquad \mathcal{H}_1 \approx \mathrm{span}\{|1\rangle\} \equiv \mathbb{C}\,,$$

so that $\mathcal{H}_0 \oplus \mathcal{H}_1 \equiv \mathbb{C} \oplus \mathbb{C}$. Here, $|0\rangle$ and $|1\rangle$ denote the states of the vacuum and the photon in the occupation number basis of the Fock space. Therefore, we have

$$L_O \cong V_O \cong A_I \cong A_O \cong B_I \cong B_O \cong A'_I \cong B'_I \cong \mathbb{C} \oplus \mathbb{C}\,.$$

Finally, the input and output spaces of beam splitters S and S' are "doubled", since a beam splitter operates over two inputs to produce two outputs. In particular,

$$\begin{array}{lllll} S_I &=& S_I^L \otimes S_I^V\,, & \qquad S'_I &=& S_I'^A \otimes S_I'^B\,, \\ S_O &=& S_O^A \otimes S_O^B\,, & \qquad S'_O &=& S_O'^A \otimes S_O'^B\,, \end{array}$$

where again

$$S_I^L \cong S_I^V \cong S_O^A \cong S_O^B \cong S_I'^A \cong S_I'^B \cong S_O'^A \cong S_O'^B \cong \mathbb{C} \oplus \mathbb{C}\,.$$

With all relevant Hilbert spaces defined, we formulate the action of each gate, using the CJ map in the form (5). The gates L and V simply generate the photon and the vacuum,

$$|L^*\rangle\!\rangle^{L_O} = |1\rangle^{L_O}\,, \qquad |V^*\rangle\!\rangle^{V_O} = |0\rangle^{V_O}\,, \tag{8}$$

where $*$ is the complex conjugation. The action of the beam splitters is

$$|S^*\rangle\!\rangle^{S_I S_O} = \left[I^{S_I S_I} \otimes (H^*)^{S_O S_I}\right] |\mathbb{1}\rangle\!\rangle^{S_I S_I}\,, \tag{9}$$

and

$$|S'^*\rangle\!\rangle^{S'_I S'_O} = \left[I^{S'_I S'_I} \otimes (H^*)^{S'_O S'_I}\right] |\mathbb{1}\rangle\!\rangle^{S'_I S'_I}\,, \tag{10}$$

where the Hadamard operator for S is defined as

$$\begin{aligned} H^{S_O S_I} = {} & \\ & \frac{1}{\sqrt{2}} \left(|1\rangle^{S_O^A} |0\rangle^{S_O^B} + |0\rangle^{S_O^A} |1\rangle^{S_O^B}\right) \langle 1|^{S_I^L} \langle 0|^{S_I^V} \\ & + \frac{1}{\sqrt{2}} \left(|1\rangle^{S_O^A} |0\rangle^{S_O^B} - |0\rangle^{S_O^A} |1\rangle^{S_O^B}\right) \langle 0|^{S_I^L} \langle 1|^{S_I^V}\,, \end{aligned}$$

and analogously for $H^{S'_O S'_I}$. The unit operator is denoted as I. Next, in the gates A and B, Alice and Bob generate their random bits a and b, and encode them into the phase of the photon. The corresponding actions are defined as

$$|A^*\rangle\!\rangle^{A_I A_O} = \left[I^{A_I A_I} \otimes (A^*)^{A_O A_I}\right] |\mathbb{1}\rangle\!\rangle^{A_I A_I}\,, \tag{11}$$

and

$$|B^*\rangle\!\rangle^{B_I B_O} = \left[I^{B_I B_I} \otimes (B^*)^{B_O B_I}\right] |\mathbb{1}\rangle\!\rangle^{B_I B_I}\,, \tag{12}$$

where

$$A^{A_O A_I} = (-1)^a |1\rangle^{A_O} \langle 1|^{A_I} \oplus |0\rangle^{A_O} \langle 0|^{A_I}\,,$$

and

$$B^{B_O B_I} = (-1)^b |1\rangle^{B_O} \langle 1|^{B_I} \oplus |0\rangle^{B_O} \langle 0|^{B_I}\,.$$

Finally, the gates A' and B' describe Alice's and Bob's measurement of the incoming state in the occupation number basis,

$$|A'^*\rangle\!\rangle^{A'_I} = |a'\rangle\,, \qquad |B'^*\rangle\!\rangle^{B'_I} = |b'\rangle\,, \tag{13}$$

where their respective measurement outcomes a' and b' take values from the set $\{0, 1\}$, depending on whether the vacuum or the photon has been measured, respectively.

After specifying the actions of the gates, the last step is the construction of the process vector

$|W_{dSD}\rangle\!\rangle$ itself. The dSD protocol assumes that the state of the photon remains unchanged during its travel between the gates. Therefore, the process vector will be a tensor product of transport vectors (3), one for each line connecting two gates in the spacetime diagram. The input and output spaces of the gates connected by the line determine the spaces of the corresponding transport vector. Thus, the total process vector is:

$$\begin{aligned}|W_{dSD}\rangle\!\rangle = \\ &|\mathbb{1}\rangle\!\rangle^{L_O S_I^L}|\mathbb{1}\rangle\!\rangle^{V_O S_I^V}|\mathbb{1}\rangle\!\rangle^{S_O^A A_I}|\mathbb{1}\rangle\!\rangle^{S_O^B B_I} \\ &|\mathbb{1}\rangle\!\rangle^{A_O S_I'^A}|\mathbb{1}\rangle\!\rangle^{B_O S_I'^B}|\mathbb{1}\rangle\!\rangle^{S_O'^A A_I'}|\mathbb{1}\rangle\!\rangle^{S_O'^B B_I'}\,.\end{aligned} \tag{14}$$

3.3 Evaluation of the probability distribution

Now that the process vector and the operations of all gates have been specified in detail, we can evaluate the probability distribution

$$p(a', b'|a, b) = \|\mathcal{M}\|^2\,, \tag{15}$$

where the probability amplitude $\mathcal{M}$ is obtained by taking the scalar product of $|W_{dSD}\rangle\!\rangle$ with the tensor product of all gates, see (7). It is most instructive to perform the computation iteratively, taking the partial scalar product of $|W_{dSD}\rangle\!\rangle$ with each gate, one by one. The explicit calculation of each step is based on two lemmas from Appendix A.

We begin by taking the partial scalar product of (14) and the preparation gates (8). Using Lemma 1 from Appendix A, we obtain:

$$\begin{aligned}\left(\langle\!\langle L^*|^{L_O} \otimes \langle\!\langle V^*|^{V_O}\right)|W_{dSD}\rangle\!\rangle = \\ &|1\rangle^{S_I^L}|0\rangle^{S_I^V}|\mathbb{1}\rangle\!\rangle^{S_O^A A_I}|\mathbb{1}\rangle\!\rangle^{S_O^B B_I} \\ &|\mathbb{1}\rangle\!\rangle^{A_O S_I'^A}|\mathbb{1}\rangle\!\rangle^{B_O S_I'^B}|\mathbb{1}\rangle\!\rangle^{S_O'^A A_I'}|\mathbb{1}\rangle\!\rangle^{S_O'^B B_I'}\,.\end{aligned}$$

Next we take the partial scalar product with the beam splitter S gate operation (9). Using Lemma 2 from Appendix A, we obtain:

$$\begin{aligned}\left(\langle\!\langle S^*|^{S_I S_O} \otimes \langle\!\langle L^*|^{L_O} \otimes \langle\!\langle V^*|^{V_O}\right)|W_{dSD}\rangle\!\rangle = \\ &\frac{1}{\sqrt{2}}\left(|1\rangle^{A_I}|0\rangle^{B_I} + |0\rangle^{A_I}|1\rangle^{B_I}\right) \\ &|\mathbb{1}\rangle\!\rangle^{A_O S_I'^A}|\mathbb{1}\rangle\!\rangle^{B_O S_I'^B}|\mathbb{1}\rangle\!\rangle^{S_O'^A A_I'}|\mathbb{1}\rangle\!\rangle^{S_O'^B B_I'}\,.\end{aligned}$$

Now we apply the Alice's gate operation (11) to obtain:

$$\begin{aligned}\Big(\langle\!\langle A^*|^{A_I A_O} \otimes \langle\!\langle S^*|^{S_I S_O} \otimes \\ \langle\!\langle L^*|^{L_O} \otimes \langle\!\langle V^*|^{V_O}\Big)|W_{dSD}\rangle\!\rangle = \\ &\frac{1}{\sqrt{2}}\left((-1)^a|1\rangle^{S_I'^A}|0\rangle^{B_I} + |0\rangle^{S_I'^A}|1\rangle^{B_I}\right) \\ &|\mathbb{1}\rangle\!\rangle^{B_O S_I'^B}|\mathbb{1}\rangle\!\rangle^{S_O'^A A_I'}|\mathbb{1}\rangle\!\rangle^{S_O'^B B_I'}\,.\end{aligned}$$

Similarly, applying Bob's gate (12) we get:

$$\begin{aligned}\Big(\langle\!\langle B^*|^{B_I B_O} \otimes \langle\!\langle A^*|^{A_I A_O} \otimes \langle\!\langle S^*|^{S_I S_O} \otimes \\ \langle\!\langle L^*|^{L_O} \otimes \langle\!\langle V^*|^{V_O}\Big)|W_{dSD}\rangle\!\rangle = \\ \frac{1}{\sqrt{2}}\left((-1)^a|1\rangle^{S_I'^A}|0\rangle^{S_I'^B} + (-1)^b|0\rangle^{S_I'^A}|1\rangle^{S_I'^B}\right) \\ |\mathbb{1}\rangle\!\rangle^{S_O'^A A_I'}|\mathbb{1}\rangle\!\rangle^{S_O'^B B_I'}\,.\end{aligned}$$

The next step is the application of the second beam splitter gate (10). After a little bit of algebra, the result is:

$$\begin{aligned}\left(\langle\!\langle S'^*|^{S_I' S_O'} \otimes \langle\!\langle B^*|^{B_I B_O} \otimes \langle\!\langle A^*|^{A_I A_O} \otimes \langle\!\langle S^*|^{S_I S_O} \otimes \langle\!\langle L^*|^{L_O} \otimes \langle\!\langle V^*|^{V_O}\right)|W_{dSD}\rangle\!\rangle = \\ \frac{(-1)^a + (-1)^b}{2}|1\rangle^{A_I'}|0\rangle^{B_I'} + \frac{(-1)^a - (-1)^b}{2}|0\rangle^{A_I'}|1\rangle^{B_I'}\,.\end{aligned}$$

Finally, applying the measurement gates (13), we obtain the complete probability amplitude,

$$\begin{aligned}\mathcal{M} \;=\; &\frac{(-1)^a + (-1)^b}{2}\delta_{a'1}\delta_{b'0} \\ &+\frac{(-1)^a - (-1)^b}{2}\delta_{a'0}\delta_{b'1}\,,\end{aligned}$$

and substituting this into (15), we obtain the desired probability distribution of the dSD process:

$$\begin{aligned}p(a', b'|a, b) \;=\; &\frac{1 + (-1)^{a+b}}{2}\delta_{a'1}\delta_{b'0} \\ &+\frac{1 - (-1)^{a+b}}{2}\delta_{a'0}\delta_{b'1}\,.\end{aligned}$$

From the probability distribution we can now conclude that there are two distinct possibilities:

either Alice detects the photon and Bob does not, $a' = 1, b' = 0$, or vice versa, $a' = 0, b' = 1$. In the first case, because total probability must be equal to one, we have

$$\frac{1 + (-1)^{a+b}}{2} = 1\,, \qquad \frac{1 - (-1)^{a+b}}{2} = 0\,.$$

The only solution to these equations is $a = b$, which means that Alice and Bob have initially generated equal bits. Since both know the probability distribution and their own bit, they both know each other's bit as well, with certainty. In the second case, when Bob detects the photon, we instead have

$$\frac{1 + (-1)^{a+b}}{2} = 0\,, \qquad \frac{1 - (-1)^{a+b}}{2} = 1\,,$$

and the only solution is $a \neq b$, meaning that Alice and Bob have initially generated opposite bits. Again, both parties know the probability distribution and their own bit, and therefore each other's bit as well, with certainty.

In order to formalise this result, one can also introduce the parity $\pi \equiv a \oplus b$ and rewrite the probability distribution in the form

$$\begin{aligned} p(a', b'|\pi) \quad = \quad & \frac{1 + (-1)^{\pi}}{2} \delta_{a'1}\delta_{b'0} \\ & + \frac{1 - (-1)^{\pi}}{2} \delta_{a'0}\delta_{b'1}\,. \end{aligned} \tag{16}$$

Thus, if Alice detects the photon, then π is even, while if Bob detects the photon, π must be odd. In both cases, they can "guess" each other's bits with certainty by calculating

$$x = \pi \oplus a\,, \qquad y = \pi \oplus b\,,$$

where x is Alice's prediction of the value of Bob's bit, and y is Bob's prediction of the value of Alice's bit. Therefore, the probability of guessing each other's input bit is

$$p_{\text{success}} \equiv p(x = b \wedge y = a) = 1\,. \tag{17}$$

3.4 Analysis of the process matrix description — operational interpretation of the vacuum

After we have given the detailed process matrix description of the dSD protocol and derived the result (17), we analyse in more detail the role of the vacuum in the formalism, giving its operational interpretation.

In order to clarify the exposition, let us give an overview of the argument, as follows:

- In the next paragraph below, we analyse the role of the vacuum in the dSD protocol, and conclude that the interaction with the vacuum should be regarded as an operation, on the same footing with all other interactions.
- In the following four paragraphs, we discuss the optical implementation of the quantum switch protocol, which also features interactions between the agents and the vacuum. Since the same physical situation should always be described in the same way, we conclude that the interaction with the vacuum should be treated as an operation in this protocol as well. Thus, the protocol features a total of four, rather than two, operations.
- Finally, in the remaining three paragraphs, we discuss the alternative point of view, namely that the interaction with the vacuum is not regarded as an operation. This is the case in the method for counting operations proposed in [22]. We conclude that it would then mean that in the dDS protocol an agent could extract information from the system at t_f without performing an operation at all.

In the dSD protocol four operations (gates), A, A', B and B' are performed (see Figure 1). Note that for each choice of input bits a and b one of the two operations performed, A' and B', is of a special form: it represents the absence of the particle. This gives rise to an *operational interpretation* of the vacuum state as a carrier of information, playing the central role in the protocol — the very interaction between the apparatus and the vacuum (the absence of a particle) plays *exactly* the same role in this protocol as any other operation, i.e., not detecting a particle ("seeing the vacuum") is an operation on its own. From the mathematical point of view, supported by the structure of the process vector (14) that explicitly features the vacuum state, it is perfectly natural to consider the interaction between the apparatus and the vacuum state on equal footing with the interaction between the apparatus and the field excitation (i.e., the particle). Both interactions equally represent operations. Therefore, one should regard the interaction with the vacuum as a resource, in the same way as the interaction with the particle.

Let us now consider the optical quantum switch, a similar protocol in which the notion of

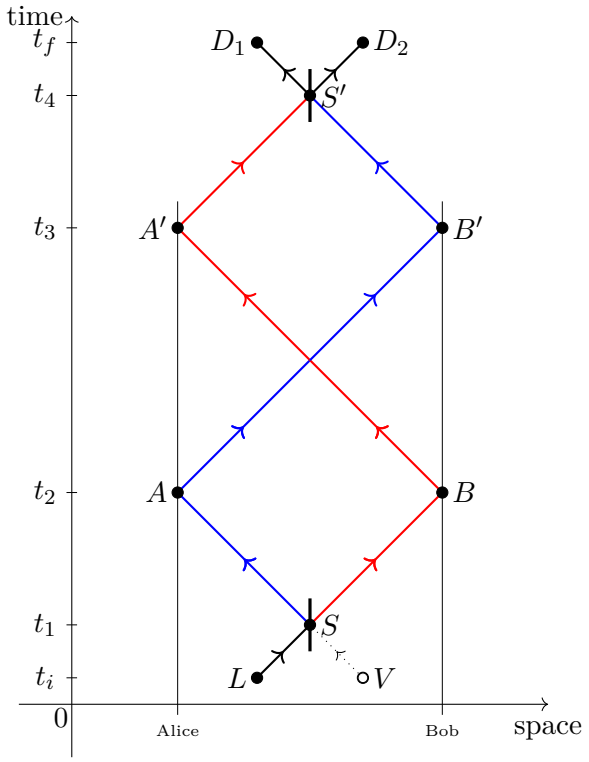


Figure 2: The complete spacetime diagram of the process corresponding to the optical quantum switch. Upon receiving the photon, Alice rotates its polarisation by the unitary U. Analogously, Bob performs rotation V on the photon entering his lab.

the vacuum also plays a role. Current optical implementations of the quantum switch feature four spacetime points, the same as the dSD protocol [9, 10, 11, 12], thus having the similar type of the spacetime schematic description, see Figure 2. However, by introducing the notion of time delocalised operations it was argued that the optical switch implements only two operations, U in spacetime points A or A', and V in spacetime points B or B' [21]. Nevertheless, the optical switch features the same apparatus-vacuum interaction as the one from the dSD protocol: whenever the particle is in, say, the blue branch, and the operations U and V are applied at spacetime points A and B', respectively, Alice's and Bob's labs experience the interaction with the vacuum at spacetime points B and A' (and analogously for the red branch). Therefore, the treatment of the vacuum in the optical quantum switch is mutually incoherent with the treatment of the vacuum in the dSD protocol.

Our analysis can thus serve as motivation for a search towards a more coherent treatment of the vacuum within the operational approach, since the same physical situation — interaction between the apparatus and the vacuum — is currently treated differently in the descriptions of the two protocols.

One might consider the following possible chain of inference. From the examples of both the quantum switch and the dSD protocol, we have that unitary operations (be it "genuine rotations" U and V, as well as phase flips $\pm I$) are considered to be operations. From the example of the dSD protocol, we see that the interaction with the vacuum is an operation as well. Further, in reference [21] it was argued that the optical switch features two "time-delocalised operations", U and V. Thus, by the same token, it follows that within this operational approach the optical switch should feature two additional "time-delocalised operations": interactions with the vacuum, one performed by Alice, and the other by Bob (see Appendix B). Therefore, the protocol features a total of four, rather than two, operations. Note that this is a possible treatment of the vacuum, which still features superposition of orders of operations U and V in the optical switch.

It is obvious that the interaction with the vacuum plays a prominent role in achieving the goal of the dSD protocol — communication between Alice and Bob. But interactions with the vacuum are also crucial in the optical switch. Indeed, without those operations, it would be *impossible* to achieve superposition of orders of operations U and V in flat spacetime with fixed causal order of spacetime points [9].

In [22] the so-called "flag" systems were introduced to count the number of operations performed in a lab without destroying the superposition, which count only one operation per each lab of the optical switch. Note though that using this method, which effectively counts the number of times a particle enters the lab, one would count three rather than four operations in the dSD protocol. This means that either the method is not appropriate, or in fact the dSD protocol features three, instead of four operations. In the case of the former, it would be useful to introduce a formal operational definition of a general method of counting operations, given that the above "flag" method cannot count interactions with the vacuum. In the case of the latter, it would mean that one could extract the information from the sys-

tem at t_f without performing an operation at all.

Indeed, if the interaction between the vacuum and the apparatus would not be considered an operation, an issue with formulating the process vector for the dSD protocol would arise. The one we formulated in (14) contains input and output Hilbert spaces associated with the interaction between the vacuum and the detectors. It is not possible to formulate a process matrix for the dSD protocol that would feature three operations, without the mentioned interaction with the vacuum. Namely, depending on the choice of input bits a and b, the photon will end up either in Alice's or Bob's lab, rendering it impossible to know in advance which of the two agents is supposed to perform the final operation. Thus, it is not possible to formulate a process matrix which features only one operation at the final moment t_f. Note that the process matrices themselves were introduced as the main tool for describing quantum processes in the operational approach. In other words, the impossibility of formulating the main operational tool for the dSD protocol without introducing the interaction with the vacuum as an operation, suggests that the latter should be considered as an operation in that protocol.

Note that, if the dSD and the optical switch protocols featured incoherent mixtures of the two possible paths instead of coherent superpositions, then one could formulate the corresponding process matrices without treating the interaction with the vacuum as an operation, indeed without even mentioning the vacuum at all. These would be purely classical processes, which would not feature any interference effects. In general, ommiting the vacuum is a natural point of view in classical physics. However, if one wants to describe quantum physics, the notions of the vacuum and its interaction with the apparatus are unavoidable.

4 Identical particles

The above analysis shows that the vacuum state plays a physically relevant role in transmitting information, and cannot be ignored. From the point of view of QFT this is a perfectly natural state of affairs, but from the point of view of quantum mechanics (QM) it is not, since the notion of vacuum as a physical state does not exist in QM a priori, and needs to be explicitly introduced by hand. Moreover, in QFT one can naturally study systems of indefinite number of identical particles. Therefore, as a first step towards the generalization of the process matrix formalism to QFT, we apply the existing abstract process matrix formalism to the representation of the second quantization.

In this section, we give basic elements of the process matrix formalism, when applied to systems of identical particles. In order to avoid working with (anti-)symmetrised vectors of multi-particle states that contain non-physical entanglement whenever two or more identical particles are fully distinguishable (say, one photon is in Alice's, and another in Bob's lab), we will use the representation of the second quantisation in which the effects of particle statistics are governed by the creation and annihilation (anti-)commutation rules. First, we need to move from the single-particle Hilbert spaces associated to the gates and the process matrix to the corresponding capped Fock spaces.

To each gate G, we assign the input/output Fock spaces, $\mathcal{G}_{I/O}$, given in terms of the vacuum state $|0\rangle$ and the single-particle Hilbert spaces $G_{I/O}$. The single-particle input Hilbert space is given as

$$G_I = \text{span}\{|i\rangle = a_i^\dagger|0\rangle \mid i = 1, 2, \dots d_I\},$$

such that its creation and annihilation operators satisfy the standard (anti-)commutation relations,

$$[a_i^\dagger, a_j^\dagger]_\pm = [a_i, a_j]_\pm = 0\,, \quad [a_i, a_j^\dagger]_\pm = \delta_{ij}\,, \tag{18}$$

where $[_\,,_]_+$ stands for anti-commutator, and $[_\,,_]_-$ for commutator. The overall bosonic input Fock space is then

$$\mathcal{G}_I = \bigoplus_{\ell=0}^{\infty} G_I(\ell), \tag{19}$$

where $G_I(0) = \text{span}\{|0\rangle\}$ is the zero-particle, $G_I(1) = G_I$ the single-particle, and

$$G_I(\ell) = \{[(a_1^\dagger)^{s_1} \dots (a_{d_I}^\dagger)^{s_{d_I}}]|0\rangle \mid s_1 + \dots + s_{d_I} = \ell\}$$

are the ℓ-particle orthogonal subspaces of the input Fock space. For fermions, each $s_i \in \{0, 1\}$, and the orthogonal sum in Equation (19) goes until d_I, instead of ∞. For a given gate, the output Fock space $\mathcal{G}_O$ is defined analogously, and we denote its creation and annihilation operators as $\tilde{a}_i^\dagger$

and $\tilde{a}_i$, respectively, in order to distinguish them from the corresponding operators in $\mathcal{G}_I$.

Our formalism is constructed for quantum circuits which consist of finite number of gates. This means that we work in the approximation of a finite number of spacetime points, as opposed to the standard QFT where one works with an uncountably infinitely many spacetime points. Thus, given the algebra (18) for the creation and annihilation operators at a single gate, the full algebra across all gates is normalised to a Kronecker delta, instead of the standard Dirac delta function. Moreover, the operators in (18) are operators in coordinate space, as opposed to the momentum space operators which are standard in QFT, since they create and annihilate modes at a given gate (i.e., a given spacetime point), instead of modes with a given momentum. Taking into account our assumption of finite number of gates, the single-particle Hilbert spaces $G_{I/O}$ are finite-dimensional, i.e., $d_{I/O} \in \mathbb{N}$. Since the gates are distinguishable, the modes assigned to different gates *always* (anti-)commute.

We restrict ourselves to the Minkowski spacetime, so that the global Poincaré symmetry implies that the vacuum state $|0\rangle$ is identical across different gates, as well as between input and output Fock spaces for a given gate. In this sense, each gate is assumed to be stationary in some inertial reference frame, since the Fock spaces of non-inertial gates would be subject to the Unruh effect. We leave the discussion of non-inertial gates and spacetimes with more general geometries for future work.

Once the Fock spaces have been defined, we pass on to the process matrix description of gate operations. Since a process matrix has to satisfy the normalisation rule (1), the corresponding input and output spaces have to be finite-dimensional. To that end, we restrict ourselves to capped Fock spaces, which contain only a finite number of elements in the sum (19), denoted $N \in \mathbb{N}$. Together with the fact that $d_{I/O}$ is finite, it follows that the capped Fock spaces are finite-dimensional. A gate operation is represented via a CJ isomorphism of the corresponding operator between the input and the output capped Fock spaces, defined in equation (2),

$$M = \Big[(\mathcal{I} \otimes \mathcal{M}) \left(|\mathbb{1}\rangle\!\rangle\langle\!\langle \mathbb{1}| \right) \Big]^T , \tag{20}$$

where the transport vector

$$|\mathbb{1}\rangle\!\rangle = \sum_{k=0}^{N} |\mathbb{1}_k\rangle\!\rangle \, , \tag{21}$$

is given in terms of k-transport vectors defined as

$$|\mathbb{1}_k\rangle\!\rangle = \sum \Big[\prod_{i=1}^{d} \frac{(a_i^\dagger)^{s_i}}{\sqrt{s_i!}} \Big] \otimes \Big[\prod_{i=1}^{d} \frac{(a_i^\dagger)^{s_i}}{\sqrt{s_i!}} \Big] |0\rangle \, , \tag{22}$$

where the sum is taken over all s_i satisfying the constraint $s_1 + ... + s_d = k$.

One special case of the general formula (20) is the case where gates destroy all coherence between k-particle sectors, for example by measuring the number of particles,

$$M = \sum_{k=0}^{N} \Big[\left(\mathcal{I} \otimes \mathcal{M}_k \right) \left(|\mathbb{1}_k\rangle\!\rangle\langle\!\langle \mathbb{1}_k| \right) \Big]^T , \tag{23}$$

where $\mathcal{M}_k$ represents the k-particle operator for the gate. The above gate represents a classical mixture of operations on each k-particle sector, as opposed to coherent superpositions of them.

Another special case of (20), which does preserve the coherence between k-particle sectors, is represented by linear operations. For a linear gate operation, one can analogously use the "vector" formalism, and the generalisation of the CJ vector (5). With a slight abuse of notation, using $\mathcal{M}$ to denote the operator instead of its superoperator, we can now write

$$\begin{aligned} |\mathcal{M}^*\rangle\!\rangle &= \Big[\mathcal{I} \otimes \mathcal{M}^* \Big] |\mathbb{1}\rangle\!\rangle \\ &= \sum_{k,k',k''=0}^{N} \Big[\mathcal{I}_k \otimes \mathcal{M}^*_{k'} \Big] |\mathbb{1}_{k''}\rangle\!\rangle \\ &= \sum_{k=0}^{N} \Big[\mathcal{I}_k \otimes \mathcal{M}^*_{k} \Big] |\mathbb{1}_{k}\rangle\!\rangle \, , \end{aligned}$$

since it is assumed that by definition

$$\Big[\mathcal{I}_k \otimes \mathcal{M}^*_{k'} \Big] |\mathbb{1}_{k''}\rangle\!\rangle \equiv 0 \, , \qquad k'' \notin \{k, k'\} \, .$$

Now, using (6) one can rewrite (20) into the form

$$\begin{aligned} M &= |\mathcal{M}^*\rangle\!\rangle\langle\!\langle \mathcal{M}^*| \\ &= \sum_{k,k'=0}^{N} \Big[\mathcal{I}_k \otimes \mathcal{M}^*_{k} \Big] |\mathbb{1}_{k}\rangle\!\rangle\langle\!\langle \mathbb{1}_{k'}| \Big[\mathcal{I}_{k'} \otimes \mathcal{M}^*_{k'} \Big]^\dagger , \end{aligned}$$

which is clearly different from the case (23), since it contains off-diagonal elements which preserve

coherence between k-particle sectors. One concrete example of this special case is the dSD protocol, discussed in the previous Section. Another example is a single-particle unitary operator

$$U = \sum_{i,j} u_{ij} \tilde{a}_i^\dagger a_j \,.$$

Then, its capped Fock-space generalisation is given as

$$\mathcal{M} = \sum_{k=0}^{N} \mathcal{M}_k = |0\rangle\langle 0| + \sum_{k=1}^{N} \frac{1}{k!} : U^{\otimes k} : ,$$

where $: U^{\otimes k} :$ is the normal ordering of $U^{\otimes k}$.

Given the capped Fock spaces and actions of instruments in all gates, a process matrix is defined in the same way as in Section 2, according to Eq. (4). A process matrix maps the tensor product of output spaces for all gates into the tensor product of input spaces for all gates. For example, if the process under consideration is a quantum circuit (see Section 2 of [9]), the corresponding process matrix can be represented as a tensor product of transport vectors, each corresponding to a wire connecting two gates. Transport vectors are defined in the same way as (21), where in (22) the first set of creation operators corresponds to the input space of the wire, while the second set corresponds to its output space. Given that a wire is connecting two gates, its input and output spaces correspond to the output and input subspaces of the two gates, respectively. A gate can in general have multiple incoming or outgoing wires attached to it. Therefore, its input (output) space is a tensor product of all output (input) spaces of the corresponding wires.

5 Conclusions

5.1 Summary of the results

In this work we have presented a detailed account of the dSD protocol, formulating it within the process matrix formalism. Analysing the role of the vacuum state in the dSD protocol and its process matrix description, we gave the operational interpretation of the vacuum. Our analysis shows that the interaction with the vacuum should be treated as an operation, on equal footing with all other interactions, thus representing a resource in quantum information protocols (including, for example, [23, 24]). As a consequence, the optical implementation of the quantum switch protocol features four rather than just two operations, in contrast to what was claimed in the literature [3, 4, 5, 6, 7, 8]. Furthermore, we have applied the process matrix formalism to the second quantisation framework restricted to capped Fock spaces, providing the description of systems of identical particles.

5.2 Discussion

The first important point of this work is the necessity of explicitly introducing the interaction with the vacuum as a legitimate operation in the dSD protocol, on equal footing with any other operation. Indeed, the very lack of detection of the particle in the protocol provides an equal amount of information as its detection (*explicit* interaction). As a consequence, instead of interpreting the absence of particle as noninteraction, one should interpret it as the interaction between the vacuum and the apparatus, and thus as an operation. Including the interaction with the vacuum as an operation poses a question of the method of counting operations in a given protocol, since the operations corresponding to the interaction with the vacuum cannot be counted.

The introduction of the vacuum into the process matrix formalism gives a natural motivation to extend the latter to the case of identical particles, both bosons and fermions, which is the second important point of this work. However, note that while employing the formalism of second quantisation, our construction still features only a discrete number of gates. This discreteness means that we still work in *particle ontology* (i.e., mechanics). Nevertheless, our construction is an important first step towards defining the process matrix formalism in *field ontology*, i.e., fully fledged QFT.

5.3 Future lines of investigation

As mentioned in the discussion, a natural next line of investigation would be a generalisation of the process matrix formalism to full, or at least perturbative, QFT. This would include an analysis of non-inertial gates and the corresponding Unruh effect. In addition, a mathematically rigorous formulation of the axioms for the process

matrix description in Fock spaces is also an important topic to be addressed. While the primary interest in process matrices lies in their application to higher order processes [25, 26], their generalisation to QFT would also be of great interest. Finally, addressing in more detail the interaction between the agent and the vacuum within the operational approach is an interesting topic of future research.

Acknowledgments

The authors wish to thank Borivoje Dakić and Flavio del Santo for useful discussions.

RF acknowledges support from DP-PMI and Fundação para a Ciência e Tecnologia (FCT) through Grant PD/BD/128636/2017, from FCT/MCTES through national funds and when applicable EU funds under the project UIDB/50008/2020, and the QuantaGENOMICS project, through the EU H2020 QuantERA II Programme, Grant Agreement No 101017733.

NP's work was partially supported by SQIG – Security and Quantum Information Group of Instituto de Telecomunicações, by Programme (COMPETE 2020) of the Portugal 2020 framework [Project Q.DOT with Nr. 039728 (POCI-01-0247-FEDER-039728)] and the Fundação para a Ciência e a Tecnologia (FCT) through national funds, by FEDER, COMPETE 2020, and by Regional Operational Program of Lisbon, under UIDB/50008/2020 (actions QuRUNNER, QUESTS), Projects QuantumMining POCI-01-0145-FEDER-031826, PREDICT PTDC/CCI-CIF/29877/2017, CERN/FIS-PAR/0023/2019, QuantumPrime PTDC/EEI-TEL/8017/2020, as well as the FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT000.

MV was supported by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia, by the bilateral scientific cooperation between Portugal and Serbia through the project "Symmetries and Quantization - 2020-2022", no. 337-00-00227/2019-09/57 supported by the Portuguese Foundation for Science and Technology (FCT), Portugal, and the Ministry of Education, Science and Technological Development of the Republic of Serbia, and by the Science Fund of the Republic of Serbia, grant 7745968, "Quantum Gravity from Higher Gauge Theory 2021" — QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

References

[1] O. Oreshkov, F. Costa, and Č. Brukner. "Quantum correlations with no causal order". Nature Communications **3**, 1092 (2012).

[2] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron. "Quantum computations without definite causal structure". Physical Review A **88**, 022318 (2013).

[3] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner. "Witnessing causal nonseparability". New Journal of Physics **17**, 102001 (2015).

[4] J. Bavaresco, M. Araújo, Č. Brukner, and M. T. Quintino. "Semi-device-independent certification of indefinite causal order". Quantum **3**, 176 (2019).

[5] P. Allard Guérin, G. Rubino, and Č. Brukner. "Communication through quantum-controlled noise". Physical Review A **99**, 062317 (2019).

[6] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther. "Experimental superposition of orders of quantum gates". Nature Communications **6**, 7913 (2015).

[7] G. Rubino, L. A Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther. "Experimental verification of an indefinite causal order". Science Advances **3**, e1602589 (2017).

[8] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M.a Zych, Č. Brukner, and P. Walther. "Experimental entanglement of temporal order". Quantum **6**, 621 (2022).

[9] N. Paunković and M. Vojinović. "Causal orders, quantum circuits and spacetime: distinguishing between definite and superposed causal orders". Quantum **4**, 275 (2020).

[10] V. Vilasini and R. Renner. "Embedding cyclic causal structures in acyclic spacetimes: no-go results for process matrices" (2022). arXiv:2203.11245.

[11] N. Ormrod, A. Vanrietvelde, and J. Barrett. "Causal structure in the presence of sectorial

constraints, with application to the quantum switch" (2022). arXiv:2204.10273.

[12] V. Vilasini. "An introduction to causality in quantum theory (and beyond)". Master's thesis. ETH. Zürich, Switzerland (2017).

[13] J. Åberg. "Operations and single-particle interferometry". Physical Review A **70**, 012103 (2004).

[14] G. Chiribella and H. Kristjánsson. "Quantum Shannon theory with superpositions of trajectories". Proceedings of the Royal Society A **475**, 20180903 (2019).

[15] H. Kristjánsson, W. Mao, and G. Chiribella. "Witnessing latent time correlations with a single quantum particle". Physical Review Research **3**, 043147 (2021).

[16] F. Del Santo and B. Dakić. "Two-way communication with a single quantum particle". Physical Review Letters **120**, 060503 (2018).

[17] F. Del Santo and B. Dakić. "Coherence equality and communication in a quantum superposition". Physical Review Letters **124**, 190501 (2020).

[18] B. Ghosh, A. Mukherjee, and S. Aravinda. "Advantage in two-way communication using non-classical states of light" (2019). arXiv:1904.04596.

[19] F. Massa, A. Moqanaki, Ä. Baumeler, F. Del Santo, J. A. Kettlewell, B. Dakić, and P. Walther. "Experimental two-way communication with one photon". Advanced Quantum Technologies **2**, 1900050 (2019).

[20] C. Portmann, C. Matt, U. Maurer, R. Renner, and B. Tackmann. "Causal boxes: Quantum information-processing systems closed under composition". IEEE Transactions on Information Theory **63**, 3277–3305 (2017).

[21] O. Oreshkov. "Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics". Quantum **3**, 206 (2019).

[22] M. Araújo, F. Costa, and Č. Brukner. "Computational advantage from quantum-controlled ordering of gates". Physical Review Letters **113**, 250402 (2014).

[23] L.-Y. Hsu, C.-Y. Lai, Y.-C. Chang, C.-M. Wu, and R.-K. Lee. "Carrying an arbitrarily large amount of information using a single quantum particle". Physical Review A **102**, 022620 (2020).

[24] Y. Zhang, X. Chen, and E. Chitambar. "Building multiple access channels with a single particle". Quantum **6**, 653 (2022).

[25] P. Perinotti. "Causal structures and the classification of higher order quantum computations". Chapter in "Time in Physics", pages 103–127. Springer International Publishing. Cham (2017).

[26] A. Bisio and P. Perinotti. "Theoretical framework for higher-order quantum theory". Proceedings of the Royal Society A **475**, 20180706 (2019).

A Two lemmas for the process matrix evaluation

Lemma 1. Let $|\Psi^*\rangle\!\rangle^{X_O} = |\Psi^*\rangle^{X_O}$ represent a gate which has no input, while it prepares the state $|\Psi\rangle \in X_O$ as its output. Then, the scalar product of that vector and the transport vector $|\mathbb{1}\rangle\!\rangle^{X_O Y_I}$ is given as:

$$ {}^{X_O}\langle\!\langle \Psi^* | \mathbb{1} \rangle\!\rangle^{X_O Y_I} = |\Psi\rangle^{Y_I} . $$

Proof. Using the fact that the transport vector is an unnormalized maximally entangled state, the explicit calculation goes as follows:

$$\begin{aligned} {}^{X_O}\langle\!\langle \Psi^* | \mathbb{1} \rangle\!\rangle^{X_O Y_I} &= \langle\Psi^*|^{X_O} \sum_k |k\rangle^{X_O} |k\rangle^{Y_I} \\ &= \sum_k \langle \Psi^* | k \rangle \, |k\rangle^{Y_I} \\ &= \sum_k \langle k | \Psi \rangle \, |k\rangle^{Y_I} \\ &= |\Psi\rangle^{Y_I} , \end{aligned}$$

where we have used the unit decomposition $I = \sum_k |k\rangle\langle k|$ and the fact that $\langle\Psi^*|k\rangle = \langle\Psi|k\rangle^* = \langle k|\Psi\rangle$.

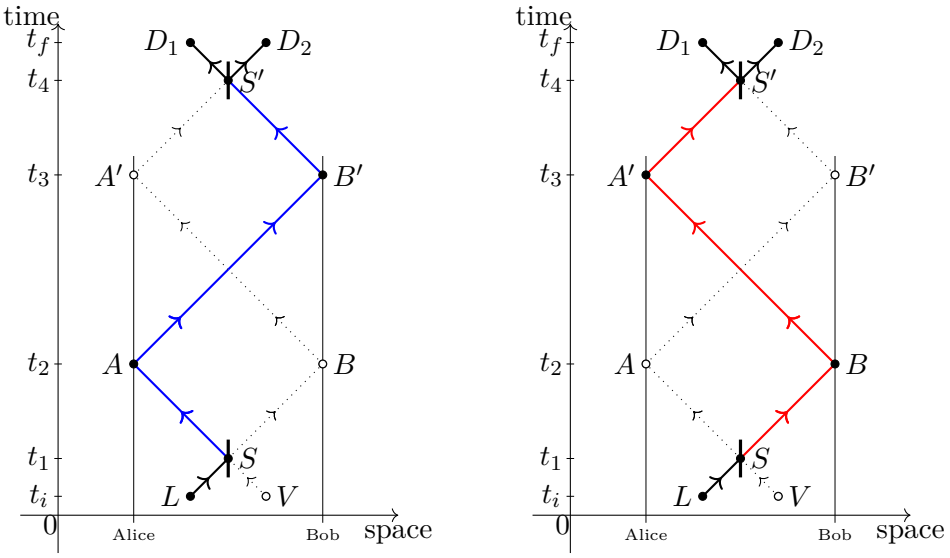


Figure 3: Two branches of the coherent superposition of the optical switch protocol.

Lemma 2. Let

$$|U^*\rangle\!\rangle^{X_I X_O} = \left[I^{X_I X_I} \otimes (U^*)^{X_O X_I}\right] |\mathbb{1}\rangle\!\rangle^{X_I X_I}$$

represent a gate which performs the operation $U : X_I \to X_O$, and let $|W\rangle\!\rangle = |\Psi\rangle^{X_I} |\mathbb{1}\rangle\!\rangle^{X_O Y_I}$. Then the scalar product of the two is

$${}^{X_I X_O}\langle\!\langle U^*|W\rangle\!\rangle = \Big(U|\Psi\rangle\Big)^{Y_I}.$$

Proof. Again using the expansion of the transport vectors as unnormalized maximally entangled states, the explicit calculation goes as follows:

$$\begin{aligned}
{}^{X_I X_O}\langle\!\langle U^*|W\rangle\!\rangle &= \langle\!\langle \mathbb{1}|^{X_I X_I} \left[I^{X_I X_I} \otimes (U^T)^{X_I X_O}\right] |\Psi\rangle^{X_I} |\mathbb{1}\rangle\!\rangle^{X_O Y_I} \\
&= \sum_k \langle k|^{X_I} \langle k|^{X_I} \left[I^{X_I X_I} \otimes (U^T)^{X_I X_O}\right] |\Psi\rangle^{X_I} \sum_m |m\rangle^{X_O} |m\rangle^{Y_I} \\
&= \sum_{k,m} \Big(\langle k|^{X_I} I^{X_I X_I} |\Psi\rangle^{X_I}\Big) \Big(\langle k|^{X_I} (U^T)^{X_I X_O} |m\rangle^{X_O}\Big) |m\rangle^{Y_I} \\
&= \sum_{k,m} \Big(\langle k|\Psi\rangle\Big) \Big(\langle m|U|k\rangle\Big) |m\rangle^{Y_I} \\
&= \sum_m \langle m|U \left(\sum_k |k\rangle\langle k|\right) |\Psi\rangle \; |m\rangle^{Y_I} \\
&= \sum_m \langle m|U|\Psi\rangle \; |m\rangle^{Y_I} \\
&= \Big(U|\Psi\rangle\Big)^{Y_I},
\end{aligned}$$

where we have again used the unit decomposition and the fact that $\langle k|U^T|m\rangle = \langle m|U|k\rangle$.

B Time-delocalised operations in the optical switch

Figure 3 depicts two branches coherently superposed in the optical switch. The left diagram represents the branch in which the photon first enters Alice's lab, and then Bob's. On the right, the photon first visits Bob's lab, and then Alice's. Whenever the photon enters Alice's lab, she applies unitary U (in A, left diagram, or A', right diagram), while Bob interacts with the vacuum (in B, left diagram, or B',

right diagram). Analogously, whenever the photon enters Bob's lab, he applies unitary V (in B, right diagram, or B', left diagram), while Alice interacts with the vacuum (in A, right diagram, or A', left diagram).

Since applying the unitaries in a quantum protocol are operations, and since in the optical switch they are applied by Alice (U) and Bob (V) at two different times, we say that the optical switch features two time-delocalised operations U (at A and A') and V (at B and B').

Since the interaction with the vacuum in the dSD protocol is an operation, and since in the optical switch it is applied by Alice and Bob at two different times, one can say that the optical switch features two time-delocalised operations of the interaction with the vacuum (at A and A', as well as at B and B').

Experimental semi-quantum key distribution with classical users

Francesco Massa[1], Preeti Yadav[2,3], Amir Moqanaki[1], Walter O. Krawec[4], Paulo Mateus[2,3], Nikola Paunković[2,3], André Souto[2,5], and Philip Walther[1]

[1]University of Vienna, Faculty of Physics, Vienna Center for Quantum Science and Technology (VCQ), Boltzmanngasse 5, Vienna A-1090, Austria

[2]Instituto de Telecomunicações, 1049-001 Lisbon, Portugal

[3]Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal

[4]Computer Science and Engineering Department, University of Connecticut, Storrs, CT 06269, USA

[5]LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal

arXiv:1908.01780v4 [quant-ph] 18 Sep 2022

Quantum key distribution, which allows two distant parties to share an unconditionally secure cryptographic key, promises to play an important role in the future of communication. For this reason such technique has attracted many theoretical and experimental efforts, thus becoming one of the most prominent quantum technologies of the last decades. The security of the key relies on quantum mechanics and therefore requires the users to be capable of performing quantum operations, such as state preparation or measurements in multiple bases. A natural question is whether and to what extent these requirements can be relaxed and the quantum capabilities of the users reduced. Here we demonstrate a novel quantum key distribution scheme, where users are fully classical. In our protocol, the quantum operations are performed by an untrusted third party acting as a server, which gives the users access to a superimposed single photon, and the key exchange is achieved via interaction-free measurements on the shared state. We also provide a full security proof of the protocol by computing the secret key rate in the realistic scenario of finite-resources, as well as practical experimental conditions of imperfect photon source and detectors. Our approach deepens the understanding of the fundamental principles underlying quantum key distribution and, at the same time, opens up new interesting possibilities for quantum cryptography networks.

1 Introduction

Quantum key distribution (QKD) is a technique that allows two distant parties, traditionally called Alice and Bob, to exchange a cryptographic key in an information-theoretic secure way. This means that the security of the key relies on information theory and cannot be broken even by an eavesdropper with unlimited resources.

The first QKD proposal was the BB84 protocol, introduced by Bennett and Brassard in 1984 [1] (subsequently, Ekert introduced the E91 protocol in 1991 [2]), which was proven secure several years later [3–5]. Since then, much progress, both theoretical and experimental, has been made in the field. The practicality of this technology is underlined by numerous experimental and even commercial endeavors, supporting its development [6–9].

Most QKD protocols require Alice or Bob to share a quantum state, or a direct quantum channel, and to perform quantum operations, i.e., operations on quantum bits (qubits) that do not have any counterpart in classical communication, such as generation or measurement in multiple bases. On the other hand, it is known that if both parties are restricted to classical communication, unconditional security is unachievable for the key distribution problem. It is therefore relevant for a fundamental understanding of QKD to investigate how quantum the users' operations and resources need to be in order to achieve information-theoretic security.

A first step in this direction was made by introducing the *semi-quantum* model of cryptography in 2007 by Boyer et al. [10]. In this model, at least one party must be "classical" in nature, i.e., restricted to a limited set of operations on qubits, namely measuring and/or preparing qubits in a single basis (usually the computational (Z) basis $\{|0\rangle, |1\rangle\}$), or simply disconnecting from the quantum channel by allowing any received quantum state to reflect back to the sender. The use of "classical" in this terminology is due to the fact that orthogonal quantum states from a single measurement basis and states of classical systems are both fully distinguishable. The other parties may be classical or quantum (naturally, at least one party must be quantum) with a "quantum" user having the ability to perform any quantum operation on qubits allowed by the laws of physics. In the subsequent proposal [11], permuting or reordering the in-

coming qubits using delay lines was considered as another classical operation. Nevertheless, although one can indeed argue that permuting physical systems is inherently classical operation, doing so, especially in photonic applications, is with the current technology far more infeasible than any quantum operation used in cryptographic protocols. Also, preparing and detecting qubit states, albeit in a single basis, is technologically non-trivial.

Further development has shown that Alice's operations can be as limited as Bob's, provided that a third party distributes entangled photons to the users and performs measurements in different bases [12, 13]. Such a scheme, referred to as a mediated SQKD protocol, allows two classical users to establish a shared secret key with one-another, using the help of a quantum server which must prepare, and later measure, quantum bits. However, this quantum server need not be trusted, and in fact could be an all-powerful adversary. Security was proven, but again, only for the perfect-qubit scenario [12].

Since that original mediated-SQKD protocol, there have been several advances both in new protocol design and in new security proof methods. A main research goal in this field is to develop new protocols which further reduce the requirements placed on either the end-users or the server (or both). In terms of reducing the complexity of the end-users, a protocol which did not require users to measure was proposed in [13] (however, attacks against the protocol were later discovered in [14]). On the other hand, in [15, 16], protocols were developed which reduced the server's requirements. Namely, in [15] the server need only send single qubit states to users but later requiring a Bell measurement. In [16] single qubits were used, both in the initial preparation stage and in the subsequent server measurement, however a cycle topology was required.

Beyond reducing end-user or server requirements, another avenue of research in this area is in improving either efficiency or noise tolerance of the protocol (or both) and in developing new security proof methods. In [17] a new multi-mediated model was introduced which could improve noise tolerance at the cost of efficiency, while in [18] a new protocol was introduced which improved efficiency (though at the cost of noise tolerance).

Most SQKD protocols up to this point have been theoretical in nature, and assume perfect qubit channels, i.e., no photon loss or multi-photons are permitted for their security to be valid. A SQKD protocol immune to such imperfections was described recently in [19] and was proven to be robust, meaning that any attack which causes an adversary to gain non-zero information on the key, necessarily creates a disturbance that may be detected with non-zero probability. A second such protocol was proposed in [20], though there security was only proven against a few specific attacks. However, no full proof of security yet exists for these protocols and so, their key rates and noise tolerances are still unknown.

In general, while numerous SQKD protocols have been proposed in the last decade [21], information-theoretic proofs of security were developed only for a few of them [12, 17, 18, 22, 23] and always in the ideal scenario of perfect qubits, ideal devices and infinite resources in the asymptotic regime.

In this work, we propose a novel SQKD protocol in the mediated model, allowing two classical users to share a secret key using the help of an untrusted, potentially adversarial, quantum server. In particular, our protocol requires Alice and Bob to perform two classical operations only, the detection or reflection of a single photon, and hence places even fewer restrictions on the users than prior protocols of this nature, by requiring only a single photon measurement and no state preparation. We are the first to show that such minimal requirements, on the part of the users, is sufficient to generate a secret key. The server's complexity is also reduced compared to prior work, needing only to prepare and measure single qubits. Furthermore, as first for mediated SQKD research, we conduct an information theoretic proof of security of the protocol assuming practical devices, whereas prior work in mediated SQKD was restricted to perfect qubit scenarios, and compute the secret key rate in the finite key setting. Finally, we experimentally demonstrate our protocol under real-life conditions and evaluate the secret key rate by using the results from actual devices. Our methods here may also be broadly applicable to other multi-user (S)QKD protocols in practical settings.

2 The Protocol

Our protocol involves three parties: two classical users, Alice and Bob, whose aim is to exchange a secret cryptographic key, and an untrusted, potentially adversarial, quantum server, which provides the quantum resources for this purpose. Furthermore, we assume that Alice and Bob can communicate through a classical authenticated channel and that the server can send unauthenticated classical messages to the users. In the description below, we discuss the protocol for single photons for simplicity, and also they are the most practical for QKD applications (though our security analysis will also take into account realistic multi-photon sources).

A sketch of the scheme is depicted in Figure 1. The server sends to Alice and Bob a single photon in a balanced superposition of their respective locations. Each user can independently choose to perform two actions: "detect" (D) or "reflect" (R). In the former case, the photon travels to a detector controlled by the user; in the latter, the photon is sent back to a balanced beam splitter controlled by the server, at

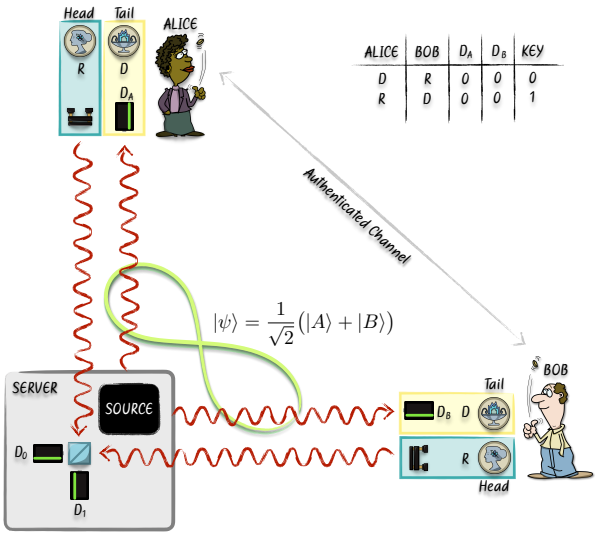


Figure 1: **The QKD protocol with classical users.** A quantum server sends single photons in superposition to the users at predetermined regular intervals, which constitute the rounds of the protocol. For each round, Alice and Bob randomly choose between "detect" (D) and "reflect" (R). The photons reflected back to the server impinge onto a beam splitter at whose outputs two detectors D0 and D1 are placed. When both Alice and Bob reflect the received photon, single-photon interference occurs at the beam splitter and only detector D0 clicks. If only one of the users chooses to detect the photon without registering any detection event, interference is suppressed and the photon has ideally 50% probability to reach detector D1. If the server announces a detection at D1 and none of the users detected photons, a raw key bit is generated according to the table in the figure. The users can communicate through a classical authenticated channel to verify the honesty of the server and to share the necessary information for the evaluation of the secure key rate.

whose output ports two detectors, D_0 and D_1, are placed. When both users choose to reflect, single-photon interference occurs at the beam splitter, with the relative phase of the two interfering photon amplitudes tuned such that only detector D_0 clicks. In the ideal case of perfect detection efficiency, when only one of the users chooses to detect the photon and does not find any, the photon collapses into the other user's location. This corresponds to performing an *interaction-free* measurement [24–26], which suppresses single-photon interference at the server and allows either detector D_0 and D_1 to click with non-zero probability. A click at detector D_1, therefore, enables each user to deduce the action of the other one, thus allowing for the establishment of a raw key digit. In particular, a key digit of 0(1) is set when Alice chooses $D(R)$ and Bob $R(D)$. Other combinations are not considered, as they cannot result in a detection at D_1. Since the raw key bits are generated when the server announces a click at detector D_1, and *neither* Alice *nor* Bob detect a photon, no use of the authenticated channels is needed during those rounds, unlike the standard QKD protocols [1, 2]. In our protocol, classically authenticated information exchange is performed only for the verification and parameter estimation rounds, which are not used to generate the raw key.

The detailed steps of the protocol are described below:

Quantum Communication Stage: Users repeat the following process until a sufficiently large raw key has been established (refer also to Figure 1):

1. The server sends a single photon to both parties in a superposition. Ideally this should be performed by the server sending a single photon through a beam splitter.

2. Alice and Bob choose, independently and randomly, between two available actions: D or R. Since Alice and Bob are completely classical, the detection results only give them information as to whether or not there is a photon at their respective detector D_A or D_B. Their actions determine their raw key bit for this round, namely:

   - **Alice:** If Alice chose D, she will record a raw key-bit of 0; otherwise, if she chose R, she will record a raw key-bit of 1.
   - **Bob:** Bob's encoding is opposite that of Alice; namely if he chose D he will record a raw key bit of 1 and, otherwise, a raw key bit of 0 if he chose R.

3. The server measures the photon coming from Alice/Bob and announces the following results: "0" if the server's detector D_0 clicks; "1" if detector D_1 clicks; "v" if no detector clicks; or "m" if more than one detector clicks. Ideally, this measurement should be performed by the server completing a (folded) Mach-Zehnder interferometer as shown in Figure 1. Note that the last case can arise due to experimental imperfections or the action of an adversary.

4. Alice and Bob perform a minimal sifting step whereby they will keep the round only if the following two conditions are met:

   - The server announces the message "1"
   - *and* Alice and Bob both did *not* detect a photon if they chose to measure.

   All other events will cause the round to be discarded. Note that, for this, Alice and Bob must announce whether they detect a photon or not. In the event parties choose R they will, by default, announce that they did not detect a photon.

Sampling Stage: Users will communicate, through an authenticated channel, their actions and measurement outcomes (if applicable) for a randomly chosen subset of the rounds performed above. This is done to verify the honesty of the server and/or the presence of an adversary. More specifically, these statistics, as

discussed below, will be used to determine a bound on the overall key-rate of the protocol.

Post Processing Stage: After performing the above sampling process and discarding those rounds chosen for sampling, users will perform a standard error correction protocol and privacy amplification protocol resulting in the final secret key of the system. For information on these standard processes, the reader is referred to [6].

It is not difficult to see that, if the server is honest, the protocol is correct. Namely, the only time the server should ever send the message "1" is when Alice and Bob choose opposite actions (thus resulting in a correlated raw key bit since their encoding operations are opposites of one another). We show later that the protocol can lead to a secure secret key even if the server is adversarial.

3 Key generation and parameter estimation

In this section, we discuss the events when raw key bits are generated and the parameter estimation procedure (for details see Appendices A, C, and D).

Let N be the total number of successful rounds in the protocol, i.e., whenever the server announces a message from "v", "0", "1" or "m". At the end of N rounds, Alice and Bob communicate with each other over an authenticated classical channel to proceed to, first, the verification procedure, and then, to estimate the parameters to eventually share a secret key among them. Note that the server is bound to announce the same results to both Alice and Bob, since it can easily be checked by the users when they communicate over an authenticated channel. Therefore, upon having all the indexed results from the server, each user compares it with their own action. During the rounds when the server announced "1", when a user either reflected, or detected vacuum, only then we say that the user's action is "consistent" with the server's result, and no information is sent to the other client. Otherwise, any of the users detecting inconsistency announces it to the other one and the corresponding round is discarded from the rounds for key-generation. Such inconsistencies could be due to receiving a click in their detectors, or receiving clicks even when they reflected due to the failure in the switch used by them to change between the actions reflect (R) or detect (D).

Therefore, when the server announces "1" and both users' actions are consistent with such outcome, then a raw key digit is generated. This occurs on total of $N_{raw} = \mathrm{p}(1)N$ rounds, where $\mathrm{p}(1)$ denotes the probability that the server announces "1" and none of the users detect any click(s). The cases when the server announces "1" and both users reflected or both detected vacuum determine errors in the key.

Note that in the majority of QKD protocols (for instance, BB84), even the very first set of keys shared by Alice and Bob requires them to communicate over an authenticated channel. On the contrary, the first set of shared key in our protocol does not require any communication between the users, but only the message "1" from the server.

Alice and Bob choose each action (R or D) independently at random, with probability 1/2. Thus, the cases when the key can potentially be generated occur with probability 1/2. In those cases, in ideal conditions, there is a probability of 1/2 that the photon collapses into the location of the user that reflects. Finally, the reflected photon has at best a further probability of 1/2 to come out from the beam splitter at the output of detector D_1. Therefore, $\mathrm{p}(1)$ is at best 1/8, which is further reduced by experimental imperfections, eavesdropping or the action of an adversarial server.

For the rest of $(1 - \mathrm{p}(1))N$ rounds, the users exchange the information of their actions and detection results over the classical channel in order to estimate the parameters necessary for the establishment of a secret key between them. Note that it is enough that only one user, say Alice, performs the verification with the information received from the other. This allows for a reduction of the communication complexity. In addition to his action choices and results for the $(1 - \mathrm{p}(1))N$ rounds, Bob can also send the messages announced by the server over all the rounds. Alice will proceed with parameter estimation only if *all* of Bob's messages match with hers.

Using the information received from Bob for the $(1 - \mathrm{p}(1))N$ rounds, Alice can perform an indirect estimation method to evaluate the probability of exchanging a key bit, p_{key}, and the probability of error on the key, p_{err}, without the need to discard any key bit. A drawback of this indirect estimation is that p_{key} and p_{err} are obtained from other directly-measured quantities, therefore, due to error propagation, their uncertainty is higher. Alternatively, the users can exchange full information about their actions for a randomly chosen fraction τ of N_{raw} rounds to directly estimate the necessary probabilities. However, in the direct estimation, the uncertainty of the final probabilities depends on the size τ of the considered sub-sample. The choice of which method to use, therefore, depends on the experimental parameters and the length of the raw key.

4 Experimental implementation

The experimental set-up for the implementation of the protocol is depicted in Figure 2. After setting its polarization to "horizontal"(H), that is parallel to the optical table, a single photon is sent to a beam splitter that creates the superposition between Alice's and Bob's locations. Each of the users controls a liquid-

crystal cell (LCC) at 45° and a polarization beam splitter (PBS). The phase retardation between the two axes of the LCC can be switched between 0 and π by means of a voltage signal. Consequently, the photon polarization is rotated by 0° or 90°, respectively. In the first case, the photon is transmitted by the PBS and steered to a fiber-coupled avalanche photo-diode (APD) for detection, D_A or D_B; in the second case, the photon travels back to the server. The detection efficiency of D_A and D_B is evaluated by comparison with a fully-characterized transition-edge superconducting nanowire detector. The photons going back to the server impinge onto a second beam splitter, at whose outputs two fiber-coupled APDs, D_0 and D_1, are placed. The set-up, therefore, implements a folded Mach-Zehnder interferometer. The phase between the two arms of the interferometer is set such that, when Alice and Bob both decide to reflect back the photon, detector D_0 clicks. The interferometer is passively stabilized, so that the phase is constant for about 100 s. After this time, the phase is actively re-set to the initial value by using a piezo transducer.

The single photons are provided by a source based on spontaneous parametric down-conversion (SPDC) in a 20 mm-long periodically-poled potassium tytanyl phosphate (PPKTP) crystal, which probabilistically converts a photon at 395 nm from a continuous-wave laser into two photons at 790 nm and with orthogonal polarizations. One photon from each produced pair is used to herald the presence of the other one, which is sent to the users. Therefore, all detections in the experiment are in coincidence with the heralding detector, D_H. The server sets intervals of 0.5 s, constituting the rounds of the protocol, in which Alice and Bob can decide to either detect or reflect the photons. Note that this interval can be made shorter, in the order of 10^{-8} s, by using ultra-fast switches and optimized bright single-photon sources [27]. At the end of each round, the server announces the result of the measurement at its detectors. The probabilistic nature of our source implies that, in each round, multiple non-simultaneous single-photon emissions can occur. In some rounds, therefore, the total number of detections is higher than one. The output rate of the source is decreased, so that the total average number of photons sent to the users is about 0.35 per round, in order to reduce the probability of multi-photon emissions.

The possibility of simultaneous multi-photon emission from the source is ruled out by the measurement of the heralded second-order correlation function at zero delay, $g^{(2)}(0)$ [28], which should be exactly 0 for an ideal perfect single-photon source. We obtain $g^{(2)}(0) = 0.004 \pm 0.010$, measured at a total detection rate of about 15×10^3 photons per round (in our case 0.5 s) and a pump power of 7 mW. Our value of $g^{(2)}(0)$ is comparable with the lowest ones obtained in quantum optics experiments [29].

5 Security analysis

We prove security of our protocol under the following assumptions:

1. The server may be compromised by the adversary. In particular, it may prepare an arbitrary initial state and perform an arbitrary quantum operation on the returning signals (both subject to the other constraints listed below). Due to this assumption, we must only analyze the case of a single adversary, namely the server, and any third party adversary's attack may be absorbed into this adversarial server's attack strategy (to the advantage of the adversary).

2. The adversary performs collective attacks only. That is, the adversary attacks by using an identical attack operation at each iteration (both for the initial state preparation strategy and the final quantum operation strategy, including the message sending). The server's initial state may be entangled with a private quantum ancilla and the final operation may also result in a private quantum memory system. The adversary is free to postpone measuring its ancilla until any future point in time and may even perform an arbitrary global measurement of its entire ancilla at that future point in time.

3. The attack performed by the adversary on each iteration of the protocol is not interactive/adaptive. In particular, the adversary must prepare an initial state once at each iteration and send it to Alice and Bob. Although this initial state may consist of multiple photons, the server cannot feed a photon into Alice or Bob's lab, and then, based on the output, immediately feed additional photons into Alice or Bob's lab. While this seems a strong assumption, there are mechanisms to enforce its compliance as we discuss in Appendix B.1. Although a full analysis of interactive attacks would be very interesting, we consider it out of scope of this paper as we are primarily focused on the development, finite key analysis, and experimental demonstration of a novel mediated SQKD protocol with minimal end-user resource requirements. We do, however, consider an interactive attack based on a "quantum bomb" attack in Appendix B.1.

4. The initial state sent by the server consists of zero, one, or two photons prepared in an arbitrary manner. This was done as our experimental implementation consisted of a negligible probability of producing three or more photons. It is also an enforceable condition if Alice and Bob used cascading interferometers to ensure the state, with high probability, does not contain

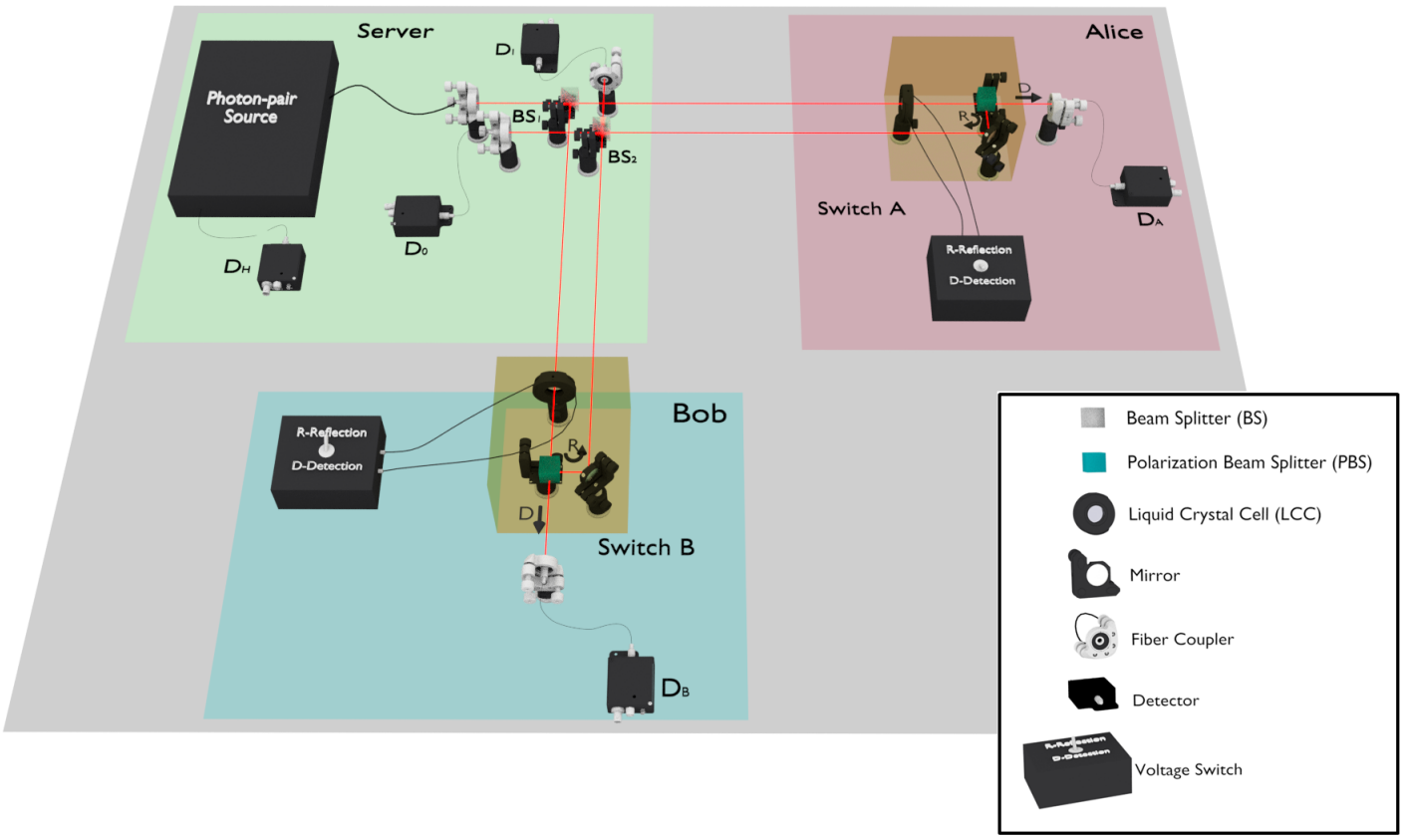


Figure 2: **Experimental set-up.** The regions of space occupied by Alice, Bob and the server are respectively marked in red, blue and green, whereas the path of the photons is indicated by red lines. The server uses a heralded single-photon source and a beam splitter (BS_1) to produce the superposition state that is sent to Alice and Bob. Each of the users controls a switch, composed of a liquid-crystal cell (LCC) at 45°, a polarization beam splitter (PBS) and a mirror. By switching the voltage of the LCC, the users can choose to steer the photon to a detector (D) or reflect it back to the server (R). The server collects the reflected photons at a second beam splitter (BS_2), where single-photon interference takes place in case both users choose to reflect. The server records the detections at D_0 and D_1 and announces the results to the users via a classical channel.

more than two photons. Our proof methodology, however, can be extended to consider the three or more photon case (assuming the attack is non-adaptive in this round as discussed above) if required. While we do not work out the exact algebra in this paper for that case, we do consider a particular multi-photon attack with three or more photons in Appendix B.2.

After Alice and Bob receive quantum states of some form from the server and perform their respective actions, they will receive a classical message from the sever indicating a possible measurement outcome. However, the server is under no obligation in our proof of security to report the measurement outcome honestly, or to even perform any measurement at all. On the rounds where the server announces "1", Alice and Bob generate the raw key of length N_{raw} whenever one of them chose to detect the photon without registering any click at the detector, while the other reflected. Note that due to experimental imperfections and eavesdropping (or server's dishonesty), server can announce "1" even if both agents reflected, or both detected vacuum, in which case they do not share the same raw key and the error is introduced. As mentioned before, from the raw key of size $N_{raw} = \mathrm{p}(1)N$, Alice and Bob may choose to use a (small) subset of size $\mu = \tau N_{raw}$ to directly estimate the statistics used to compute the secret key rate. The portion of the raw key remaining after parameter estimation step is called the sifted key, of the length $N_{sift} = N_{raw} - \mu$. Let the random variables $\mathcal{R}_A$ and $\mathcal{R}_B$ denote Alice's and Bob's respective sifted keys. After the quantum communication and sampling stages, it is not necessarily true that $\mathcal{R}_A$ and $\mathcal{R}_B$ are uniformly distributed or fully correlated. It is also not necessarily true that they are completely secret. Thus, the protocol must perform a classical post processing stage which further processes these raw key strings through error correction (to ensure they are perfectly correlated with high probability) and privacy amplification (which ensures that Eve's ancilla is independent of the final secret key.

The security level of the key shared between Alice and Bob is given by parameter ϵ, which quantifies how

uncorrelated the secret key is from Eve or a dishonest server. More formally, from [30, 31], one should have:

$$\left|\left|\rho_{KE} - \frac{I_K}{2^\ell} \otimes \rho_E\right|\right| \leq \epsilon, \tag{1}$$

where ρ_{KE} is the classical-quantum state modeling the secret key (after error-correction and privacy amplification) and Eve's ancilla, while $I_K/2^\ell \otimes \rho_E$ is an ideal uniform random key of size ℓ-bits independent of Eve. The security criterion requires ϵ to tend to zero as the number of rounds N tends to infinity, thus obtaining perfectly secret key in the asymptotic scenario. One can compute the sifted key rate as $r' = \lim_{N\to\infty} \ell/N_{sif} = S(A|C) - H(A|B)$ using results in [5]. Conditional Shannon entropy $H(A|B)$ can be easily computed using the probabilities $p_{i,j}$ of Alice and Bob establishing the raw key bit values i and j, respectively. Further, the secret key rate is defined as $r = \ell/N = r'(N_{sift}/N)$, which is the same as the sifted key rate in the asymptotic regime: since in order to obtain good enough statistics during the verification procedure, the number μ, albeit big, is still finite, we have $N_{sif} = N - \mu \approx N$, for $N \to \infty$. In the realistic case of limited resources, however, where Alice and Bob can exchange only a finite number of keys, we must take into account the imperfect parameters. Using the security criterion given by [31], let us denote ϵ_{PE} as a given error tolerance for the parameter estimation. One can further compute δ, as a function of ϵ_{PE}, a confidence interval so that the observed parameters are δ close to the actual values, except with probability ϵ_{PE}. Let ϵ be the desired security of the final secret key, and let ϵ_{EC} be the maximal probability that Bob computes error correction incorrectly. All of these are given by the user. Therefore, after μ rounds are used for the direct method of parameter estimation, the proportion of qubits used for estimating the secret key rate is $(\mathrm{p}(1)N - \mu)/N$. Using the results shown in [31], under the assumption of collective attacks, we have the following Theorem:

Theorem 1. *(Modified from [31]): Let $\rho_{AC}^{\otimes N}$ be the state of the quantum system produced by executing the protocol N times. Then, the key-rate r is bounded by:*

$$r \geq \frac{p(1)N - \mu}{N}\left(S(A|C)_\rho - \frac{\texttt{leak}_{EC} + \Delta}{p(1)N - \mu}\right), \tag{2}$$

where

$$\Delta = 2\log_2\left(\frac{1}{2(\epsilon - \epsilon_{EC} - \epsilon')}\right) + 7\sqrt{(p(1)N - \mu)\log_2(2/(\epsilon' - \epsilon_{PE}))}. \tag{3}$$

Above, $S(A|C)_\rho$ is the conditional von Neumann entropy of Alice's raw key bit register conditioned on the server's quantum memory system. The value $\texttt{leak}_{EC}$ quantifies the error-correction leakage (namely, the number of classical bits exchanged between Alice and Bob during the error correction protocol). Finally, ϵ is the desired distance to an ideal key (as in Equation 1); ϵ_{PE} is the user specified error tolerance for the parameter estimation; ϵ_{EC} is the failure probability of the error correction protocol; and ϵ' is arbitrary (chosen by the user to maximize the expression) but bounded by $\epsilon - \epsilon_{EC} > \epsilon' > \epsilon_{PE} \geq 0$.

Of course, users don't have an exact description of ρ needed to directly compute $S(A|C)$ above. Thus, to actually compute the key-rate r, $S(A|C)$ is minimized over all observable statistics within the given confidence interval (so that the actual statistics of the real density operator are within $\delta(\epsilon_{PE})$ of the observed statistics, except with probability ϵ_{PE}). Later, in our security proof, we will use a theorem from [32], stated below as Theorem 2, to actually bound the entropy $S(A|C)$. The value $\texttt{leak}_{EC}$ represents the number of (classical) bits exchanged between Alice and Bob during the error correction. Again, using [31], we take $\texttt{leak}_{EC}/(\mathrm{p}(1)N - \mu) = (1.2)h(Q)$, where $Q = \mathrm{p}_{err}/\mathrm{p}(1)$ and p_{err} is the probability to generate opposite key bits during the entire protocol. Note that μ will also be a function of ϵ_{PE}, since the smaller that is, the larger μ will be.

Theorem 2. *(From [32]): Let ρ_{AC} be a quantum state of the form:*

$$\frac{1}{N}\sum_{a=0}^{2} |a\rangle\langle a|_A \otimes \left(\sum_{i=1}^{N} |F_i^a\rangle\langle F_i^a|_C\right). \tag{4}$$

Then, the von Neumann entropy $S(A|C)_\rho$ may be bounded by

$$S(A|C)_\rho \geq \frac{1}{N}\sum_{i=1,N}\left(\langle F_i^0|F_i^0\rangle + \langle F_i^1|F_i^1\rangle\right) \times \left[h\left(\frac{\langle F_i^0|F_i^0\rangle}{\langle F_i^0|F_i^0\rangle + \langle F_i^1|F_i^1\rangle}\right) - h(\lambda_i)\right],$$

where

$$\lambda_i = \frac{1}{2}\left(1 + \frac{\sqrt{(\langle F_i^0|F_i^0\rangle - \langle F_i^1|F_i^1\rangle)^2 + 4\mathrm{Re}^2\langle F_i^0|F_i^1\rangle}}{\langle F_i^0|F_i^0\rangle + \langle F_i^1|F_i^1\rangle}\right).$$

At a high level, our security proof involves bounding the conditional von Neumann entropy $S(A|C)$ of the system assuming an adversarial server. This is achieved by first writing out an explicit description of the overall state's density operator (including the photons in the interferometer, the agents, and the Server/adversary). We then show how certain important qualities of the state, namely the overlap of various ancilla vectors of the adversary, may be determined through observable statistics (such as, for instance, p_{err}). Finally, we use Theorem 2 to bound the conditional entropy and Theorem 1 to determine a final bound on the secret key rate. These steps are algebraically involved and so are derived in detail in the appendices. Namely, in Appendix C we derive the key rate for the ideal-qubit case. This first stage also

helps to develop the intuition of the proof used for the more complicated scenario involving practical device imperfections, presented in Appendix D. Bounding $S(A|C)$ is the critical, and challenging, element of any QKD security proof. The techniques to bound this quantity developed in this work may be useful in other protocols as well.

Our security analysis takes into account the finite detection efficiencies of commercial single-photon detectors and multi-photon components in the quantum state received by Alice and Bob (see Appendix A), but does not consider other imperfections which can be used by an eavesdropper to gain information about the key. This is in general an issue for all cryptographic protocols, both classical and quantum, as it is in practice very challenging to consider all potential side channels in the security analysis [33–37]. However, specific attacks can be countered by technical adaptations of the experimental set-up. As an example, let us consider the frequency dependence of the APD's detection efficiency. By sending photons at frequencies outside the detection bandwidth of the users' detectors, an eavesdropper can in fact gain information about the agents' actions while remaining completely undetected. This specific issue can be solved by employing bandpass filters that block any incoming light at undetectable frequencies. Similar strategies can be used for other degrees of freedom which the eavesdropper could exploit to prepare undetectable photons (e.g. time, spatial mode, etc.). Current photonic technology provides effective filtering systems for all these degrees of freedom [38–41], which allows the users to counter the described category of attacks at the price of a more complicated set-up and a reduction in the secret key rate.

6 Experimental Results

To obtain the numerical values from the lower-bounds on $S(A|C)$ and other terms from the expression (2) for the secret key rate, r, we measure the probability of the raw key generation, p_{key}, and the probability of error in the raw key, p_{err}, after 10^5 rounds of the protocol. Formally, p_{key} is defined to be the probability of Alice and Bob not rejecting a round, while p_{err} is the probability that, conditioned on a raw key bit being distilled, that the raw key bit contains an error (e.g., Alice has a 0 while Bob has a 1). Note that 10^5 rounds is not sufficient to actually produce a secret key through this protocol under these operating conditions as our later evaluations show; however, it is sufficient as a proof of concept to gather experimental statistics and evaluate what the key-rate *would be* had we continued the experiment for a longer duration.

The values of p_{key} and p_{err} are evaluated in three different ways: direct estimation over the full data set, direct estimation over a randomly chosen subset of 10^4 rounds and indirect estimation. In the direct estimation, the users sacrifice a part of the raw key for verification procedure (see Appendix E.1 for details). In the indirect estimation, discussed in detail in Appendix E.2, Alice obtains p_{key} and p_{err}, using the information received from Bob during the verification phase. This allows the parties to avoid the loss of key digits, at a price of higher uncertainty on the estimated values, which are calculated from several experimentally obtained quantities, each with its error. The results are reported in Table 1.

	Direct Method (full dataset)	Direct Method (subset)	Indirect Method (full dataset)
p_{key}	$1.55(3) \times 10^{-2}$	$1.5(1) \times 10^{-2}$	$1.5(3) \times 10^{-2}$
p_{err}	$7.5(8) \times 10^{-4}$	$5(2) \times 10^{-4}$	$3(3) \times 10^{-3}$

Table 1: **Evaluation of key generation and error rates**. The probabilities of raw-key generation, p_{key} and error on a key digit, p_{err}, respectively, are shown per round (in our case an interval of 0.5 s). In the table, the numbers in parentheses are the errors on the last digits, obtained with the assumption of poissonian uncertainty on the counts.

Based on the probabilities in Table 1, we obtain the dependence of the final secret key rate, r, on the number of rounds, N, see Equation (2). This dependence is plotted in Figure 3, for different values of the detection losses of $\mathrm{D_A}$ and $\mathrm{D_B}$, assumed to be the same. The details of how the curves were obtained are discussed in Appendices D and F. As expected, an increase in the detection loss degrades the performance of the protocol.

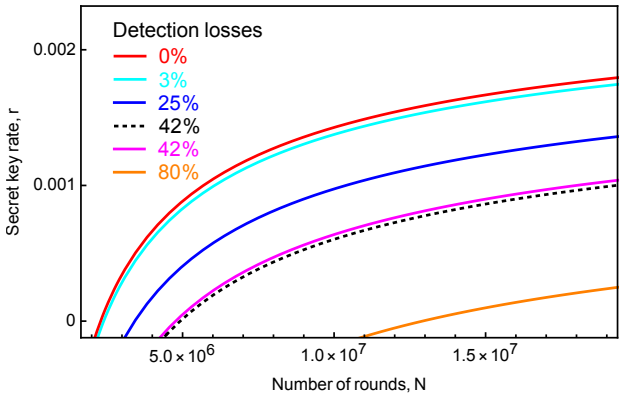


Figure 3: **Secret key rate vs number of rounds, for different values of detection loss**. The black dashed curve refers to the experimental implementation, corresponding to a detection loss of 42% for each Alice and Bob. The red, cyan, blue, magenta and orange curves represent the calculated results for detection losses of $0, 3, 25, 42$ and 80%, respectively. If the detection loss increases, the number of rounds for which r becomes positive also increases, while the asymptotic secret key rate decreases. In the implemented case, the secret key rate becomes positive after about 4.9×10^6 rounds.

We also report in Figure 4 the dependence of the secret key rate on the loss in the quantum channel between the server and each user, assumed to be the same for both, Alice and Bob. We present plots for

different values of the detection efficiency and the quantum bit error rate (QBER), which is defined as the fraction of errors in the sifted key. More details on how these plots are obtained can be found in Appendix G.

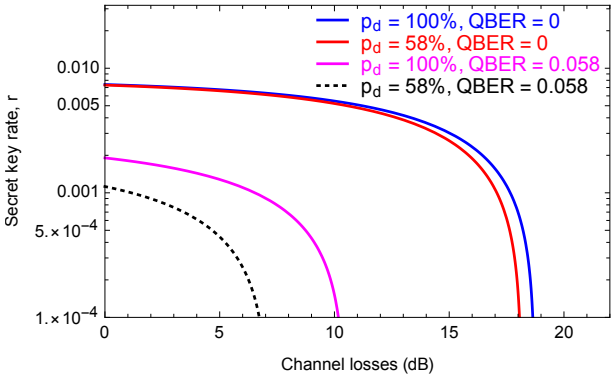


Figure 4: **Secret key rate vs channel transmission loss, for different values of detection efficiency and quantum bit error rate (QBER)**. The black dashed curve corresponds to the experimental parameters, with a detection efficiency of 58% for each Alice and Bob and a QBER of 0.058. The magenta, red and blue curve represent the three following ideal situations, respectively: perfect detectors and experimentally obtained QBER, imperfect detectors and QBER = 0, perfect detectors and QBER = 0. All curves are obtained by considering 10^9 rounds of the protocol and the statistics of the used single-photon source.

Given the results of Figure 4, we can compare the performance of our protocol to that of other QKD schemes. A natural candidate for the comparison is measurement-device-independent (MDI) QKD [42], which also involves an external server performing the detection. To our knowledge the best implementation to date of MDI-QKD achieves a secret key rate of about 10^{-4} for 7 dB of channel transmission loss [43]. We obtain a similar key rate at the same transmission loss, as shown by the dashed line in Figure 4. However, the secret key rate for our experimental parameters quickly decreases for higher losses, contrary to the realization in [43], where a secret key rate of 4.9×10^{-6} is reported for 20.4 dB of loss. Nevertheless, by considering QBER = 0, we obtain rates of the order of 10^{-4} for about 18 dB of channel loss. These results indicate that our protocol can perform as good as MDI-QKD for transmission losses up to about 7 dB, at least within the boundaries of our experimental implementation. At the moment, it is not clear if the performance could be made comparable also for higher losses, which however would require a more advanced experimental realization of our protocol.

Additionally, we stress that our estimated rates are lower-bounds and the actual key rates could be significantly higher. Indeed, to compute these lower bounds on $S(A|C)$, we took advantage of the strong subadditivity of von Neumann entropy by actually discarding several components of the entropy function (components which would only have increased Eve's uncertainty – thus, by discarding them, we are giving an unrealistic advantage to the adversary causing the key rate to drop). Such a method gives a worst-case computation.

7 Conclusions

In our work, we propose and experimentally implement a novel QKD protocol allowing two classical users to establish a shared secret key using the services of an untrusted quantum server, which provides a superimposed single photon as a feasible quantum resource. We underline the applicability of our scheme by providing an information-theoretic security analysis of our protocol in the finite-key setting, which takes into account imperfect detection efficiency and multi-photon emission from the source, and by calculating the secret key rate.

Experimentally, the main challenge of the protocol is that it requires phase stability in the interferometer formed between the users and the server. This issue can be addressed by using intrinsically phase-stable schemes, like Sagnac configurations [44]. In this case, however, a quantum channel between Alice and Bob is also necessary.

As an immediate future line of research, our security analysis of finite keys in the presence of experimental imperfections can be applied to show the same security levels for other cryptographic schemes, such as counterfactual quantum cryptography [45–49], or the key distribution based upon recently proposed two-way communication with one photon [50, 51].

In practical terms, recent progresses in bright deterministic single-photon sources [52], high-efficiency detectors [53] and fast switches [27] promise to push our scheme towards real-world applications.

Acknowledgments

We would like to thank Giulia Rubino for help with some figures and Borivoje Dakić and Ämin Baumeler for useful discussions. P.Y., P.M., N.P. and A.S. acknowledge the support of SQIG – Security and Quantum Information Group, the Instituto de Telecomunicações (IT) Research Unit, ref. UIDB/50008/2020 (actions QuRUNNER, QUESTS), funded by Fundação para a Ciência e Tecnologia (FCT), and the FCT projects QuantumMining POCI-01-0145-FEDER-031826, Predict PTDC/CCI-CIF/29877/2017 and QuantumPrime PTDC/EEI-TEL/8017/2020, supported by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme (COMPETE 2020) of the Portugal 2020 framework [Project Q.DOT with Nr. 039728 (POCI-01-0247-FEDER-039728)], and by the Regional Operational Program of Lis-

bon. P.Y. acknowledges the support of DP-PMI and FCT (Portugal) through the scholarship PD/BD/113648/2015. W.K. is partially supported by NSF Grant No. 1812070. N.P. acknowledges FCT project CERN/FIS-PAR/0023/2019, as well as the FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT000. A.S. acknowledges funds granted to LaSIGE Research Unit, ref. UID/CEC/00408/2013. P.W. acknowledges support from the research platform TURIS, from the European Commission through ErBeStA (No.800942), from the Austrian Science Fund (FWF) through BeyondC (F7113-N48) and Research Group 5 (FG5), and from the U.S. Air Force Office of Scientific Research (FA9550-21-1-0355) **Author contributions:** P.Y., W.K., P.M., N.P., A.S. developed the protocol and analysed its security. F.M., A.M. and P.W. designed and implemented the experiment and analysed the experimental data. All the authors contributed to the writing of the final manuscript. F.M. and P.Y. contributed equally to this work. **Corresponding authors:** correspondence to Francesco Massa (francesco.massa@univie.ac.at) or Philip Walther (philip.walther@univie.ac.at).

A Extraction of the secret key

In order to compute the secret key rate described above, one needs to compute $S(A|C)$ for a given system. Before we proceed to discuss the ideal and experimental scenario, let us first define some useful terminology.

Let us denote the Hilbert spaces corresponding to Alice's and Bob's equipments as $\mathcal{H}_A = \text{span}\{|D_c\rangle_A, |D_v\rangle_A, |D_\ell\rangle_A, |D'_\ell\rangle_A, |D'_c\rangle_A, |R\rangle_A\}$ and $\mathcal{H}_B = \text{span}\{|D_c\rangle_B, |D_v\rangle_B, |D_\ell\rangle_B, |D'_\ell\rangle_B, |D'_c\rangle_B, |R\rangle_B\}$, respectively. Here, $|D_c\rangle$ and $|D_v\rangle$ denote the states of a detector, the first corresponding to the case of a photon causing a click, and the second corresponding to the case when there were no photons, resulting in a no-click. The detectors' state corresponding to the case when an incoming photon was lost is denoted as $|D_\ell\rangle$. The state $|D'_\ell\rangle$ corresponds to a loss, while $|D'_c\rangle$ to a click, of the photon at time $t' \neq t$, when two non-simultaneous photons were emitted by the source at times t and t'. Finally, $|R\rangle$ denotes the state of a reflecting mirror. Note that the states corresponding to a click, $|D_c\rangle$ and $|D'_c\rangle$, and the ones corresponding to no-click, $|D_v\rangle$, $|D_\ell\rangle$ and $|D'_\ell\rangle$ are macroscopically distinguishable between each other as groups of those with or without clicks; and also to $|R\rangle$. However, the first two, $|D_c\rangle$ and $|D'_c\rangle$, are not distinguishable among each other, since in our set-up, Alice and Bob do not keep track of the detection times. Moreover, the latter three states, $|D_v\rangle$, $|D_\ell\rangle$ and $|D'_\ell\rangle$, also cannot be distinguished among each other, since without performing sophisticated quantum measurements, one cannot distinguish whether a detector did not click because there were no photons present, or they were lost.

We denote the server's Hilbert space as $\mathcal{H}_S = \text{span}\{|0\rangle_S, |1\rangle_S, |v\rangle_S, |m\rangle_S\}$ consists of macroscopic orthogonal states modeling classical messages "0", "1", "v" (vacuum) and "m" (multiple clicks), respectively. Additionally, we denote server's ancilla system by C, spanned by the Hilbert space $\mathcal{H}_C$, which a dishonest server can entangle with the photons sent to Alice and Bob to extract information about the exchanged key.

Let us assume Alice tosses a fair coin to decide whether she will detect or reflect the photon, and set the initial state of the apparatus accordingly, resulting in a proper mixture of the two states, $|D_v\rangle_A \langle D_v|$ and $|R\rangle_A \langle R|$, and analogously for Bob. Without the loss of generality, we can always include the coin states into the macroscopic description of the apparatus states, such that the purified initial state of Alice's apparatus is

$$|\phi_0\rangle_A = \frac{1}{\sqrt{2}}\Big(|D_v\rangle_A + |R\rangle_A \Big), \tag{5}$$

and analogously for Bob, making their joint state as

$$|\phi_0\rangle_{AB} = \frac{1}{2}\Big(|D_v,R\rangle_{AB} + |R,D_v\rangle_{AB} + |D_v,D_v\rangle_{AB} + |R,R\rangle_{AB}\Big). \tag{6}$$

Note that due to possible imperfect single-photon sources, and the presence of adversaries, the number of photons present is not necessarily fixed to be one. Thus, we will use a number basis to describe the photonic states. In this paper, we will decompose the overall Fock space of the photons in Alice's and Bob's arms as $\mathcal{F}_f = \text{span}\{|0,0\rangle_f, |1,0\rangle_f, |0,1\rangle_f, |2,0\rangle_f, |1,1'\rangle_f, |1',1\rangle_f, |0,2\rangle_f\} \oplus \mathcal{F}_f^k$, where $|0,0\rangle_f \equiv |v\rangle_f$ represents the vacuum state, $|1,0\rangle_f$ represents a photon in Alice's arm and $|0,1\rangle_f$ to be in Bob's arm. Similarly, $|2,0\rangle_f$, and $|0,2\rangle_f$, represent two non-simultaneous photons in Alice' and Bob's arms, respectively; whereas $|1,1'\rangle_f$ and $|1',1\rangle_f$ represent the case of two non-simultaneous photons when the first one went to Alice's arm while the second to Bob and vice-versa, respectively. $\mathcal{F}_f^k$ denotes the sub-space corresponding to the multi-photon case of $k > 2$ photons. The action of photonic creation operators $\hat{a}^\dagger$ and $\hat{b}^\dagger$, in terms of the number basis $|a,b\rangle_f$, with $a, b \in \mathbb{N}_0$ being the number of photons in Alice's and Bob's arms, respectively, is given by $(\hat{a}^\dagger)^a(\hat{b}^\dagger)^b |0\rangle_f = \sqrt{a!\,b!}\,|a,b\rangle_f$. We can now proceed to analyze the experimental implementation of our protocol with imperfect single-photon sources and detectors, as well as the noisy and lossy channels.

We assume an untrusted server that can attack before Alice and Bob perform their respective operations, as well as after (which is equivalent to allowing Eve to intercept the photons exchanged between an honest server and the agents). We consider a poissonian probabilistic single photon source, emitting vacuum state with probability p_0, single photons with probability p_1, two non-simultaneous photons with probability p_2, etc., within a time slot of interval T, as

$$|\phi_0\rangle_f = \sqrt{\mathrm{p}_0}\,|v\rangle_f + \sqrt{\frac{\mathrm{p}_1}{T}}\int_0^T \hat{a}^\dagger(t)\,|v\rangle_f\,\mathrm{d}t + \frac{\sqrt{\mathrm{p}_2}}{T}\int_0^T\int_0^T\left(\frac{\hat{a}^\dagger(t)\hat{a}^\dagger(t')}{\sqrt{2}}\,|v\rangle_f\right)\mathrm{d}t\,\mathrm{d}t' + \ldots, \tag{7}$$

where $\hat{a}^\dagger(t)$ and $\hat{a}^\dagger(t')$ represent photon creation at times t and t', respectively. In our particular implementation, the average number of photons is 0.35, yielding $\mathrm{p}_0 = 0.705$, $\mathrm{p}_1 = 0.247$, $\mathrm{p}_2 = 0.043$. For simplicity, and in order to compare the theoretical analysis with our experimental data, the probability to emit higher numbers of photons is considered negligible, i.e., $\mathrm{p}_0 + \mathrm{p}_1 + \mathrm{p}_2 \approx 1$. Thus, the initial photon state is

$$|\phi_0\rangle_f = \sqrt{\mathrm{p}_0}\,|v\rangle_f + \sqrt{\mathrm{p}_1}\,|1\rangle_f + \sqrt{\mathrm{p}_2}\,|2\rangle_f\,, \tag{8}$$

where $|v\rangle_f \equiv |0\rangle_f$ is the photon vacuum state, $|1\rangle_f = \hat{a}^\dagger(t)\,|v\rangle_f$, $\sqrt{2}\,|2\rangle_f = \hat{a}^\dagger(t)\hat{a}^\dagger(t')\,|v\rangle_f$. Nevertheless, our analysis can straightforwardly generalised to an arbitrary number of emitted photons. Note that, for simplicity, we omitted the time integrals in the definition of the single- and two-photon states, $|1\rangle_f$ and $|2\rangle_f$, respectively, as we consider that the users do not keep track of the photon detection times, meaning that, at the end of each round, Alice, Bob and the server only have access to the number of detections they recorded. This makes our analysis also applicable to the case of simultaneous multi-photon emission.

After passing through the first 50/50 beam splitter of our interferometer, described by $\hat{a}^\dagger(t) \to (\hat{a}^\dagger(t) + \hat{b}^\dagger(t))/\sqrt{2}$ and $\hat{a}^\dagger(t') \to (\hat{a}^\dagger(t') + \hat{b}^\dagger(t'))/\sqrt{2}$, the above state becomes

$$|\phi_0\rangle_f = \sqrt{\mathrm{p}_0}\,|v\rangle_f + \sqrt{\frac{\mathrm{p}_1}{2}}\Big(\,|1,0\rangle_f + |0,1\rangle_f\,\Big) + \frac{\sqrt{\mathrm{p}_2}}{2}\Big(\,|2,0\rangle_f + |1,1'\rangle_f + |1',1\rangle_f + |0,2\rangle_f\,\Big). \tag{9}$$

Upon possible further action of the adversary, the most general photon-server (normalized) state is given by

$$\begin{aligned}
|\phi_0\rangle_{fC} &= \sum_{\substack{a,b\geq 0\\ a+b\leq 2}} |a,b\rangle_f\,|c_{a,b}\rangle_C \\
&= |0,0\rangle_f \otimes |c_{0,0}\rangle_C \\
&\quad + |1,0\rangle_f \otimes |c_{1,0}\rangle_C + |0,1\rangle_f \otimes |c_{0,1}\rangle_C \\
&\quad + |2,0\rangle_f \otimes |c_{2,0}\rangle_C + |0,2\rangle_f \otimes |c_{0,2}\rangle_C + |1,1'\rangle_f \otimes |c_{1,1'}\rangle_C + |1',1\rangle_f \otimes |c_{1',1}\rangle_C\,.
\end{aligned} \tag{10}$$

where $|c_{a,b}\rangle_C \in \mathcal{H}_C$ (not necessarily orthogonal, nor normalized states) are associated to the cases when there are a and b photons entering Alice's and Bob's arms, respectively. Nevertheless, the states $|c_{a,b}\rangle_C$ are arbitrary and contain any number of photons. Therefore, the overall state before the photon(s) enter Alice's and Bob's labs is

$$\begin{aligned}
|\phi_0\rangle_{ABfC} &= |\phi_0\rangle_{AB} \otimes |\phi_0\rangle_{fC} \\
&= \frac{1}{2}\Big(|D_v,D_v\rangle_{AB} + |D_v,R\rangle_{AB} + |R,D_v\rangle_{AB} + |R,R\rangle_{AB}\Big) \\
&\quad \otimes \Big(|0,0\rangle_f \otimes |c_{0,0}\rangle_C \\
&\qquad + |1,0\rangle_f \otimes |c_{1,0}\rangle_C + |0,1\rangle_f \otimes |c_{0,1}\rangle_C \\
&\qquad + |2,0\rangle_f \otimes |c_{2,0}\rangle_C + |0,2\rangle_f \otimes |c_{0,2}\rangle_C + |1,1'\rangle_f \otimes |c_{1,1'}\rangle_C + |1',1\rangle_f \otimes |c_{1',1}\rangle_C\Big).
\end{aligned} \tag{11}$$

Let us denote Alice's and Bob's respective detectors' efficiencies as p_d^A and p_d^B, with the respective losses being $\mathrm{p}_\ell^A = 1 - \mathrm{p}_d^A$ and $\mathrm{p}_\ell^B = 1 - \mathrm{p}_d^B$. In our experimental implementation, the two efficiencies are almost the

same, with $\mathrm{p}_d^A \approx \mathrm{p}_d^B \approx 58\%$. The individual actions of, say, Alice, in this practical scenario are

$$
\begin{aligned}
&|D_v\rangle_A |0\rangle_f \to |D_v\rangle_A |0\rangle_f ,\\
&|D_v\rangle_A |1\rangle_f \to \left(\sqrt{\mathrm{p}_\ell^A}\,|D_\ell\rangle_A + \sqrt{\mathrm{p}_d^A}\,|D_c\rangle_A\right)|0\rangle_f ,\\
&|D_v\rangle_A |2\rangle_f \to \Big(\mathrm{p}_\ell^A\,|D_\ell D'_\ell\rangle_A + \sqrt{\mathrm{p}_\ell^A \mathrm{p}_d^A}\,|D_c D'_\ell\rangle_A\\
&\qquad\qquad +\sqrt{\mathrm{p}_\ell^A \mathrm{p}_d^A}\,|D_\ell D'_c\rangle_A + \mathrm{p}_d^A\,|D_c D'_c\rangle_A\Big)|0\rangle_f ,\\
&|R\rangle_A |0\rangle_f \to |R\rangle_A |0\rangle_f ,\\
&|R\rangle_A |1\rangle_f \to |R\rangle_A |1\rangle_f ,\\
&|R\rangle_A |2\rangle_f \to |R\rangle_A |2\rangle_f ,
\end{aligned}
\tag{12}
$$

where primed and unprimed states of the apparatuses correspond to at times t' and t, respectively. Note that we assume that Alice and Bob trust their detectors with their finite detection efficiencies. Therefore, upon applying U_1, given in terms of Alice's and Bob's local actions described by (12), we obtain the state $|\phi_1\rangle_{ABfC} = U_1\,|\phi_0\rangle_{ABfC}$.

Upon leaving Alice's and Bob's labs, the server (or Eve) will apply a quantum instrument to the returning photon-server state. This can be modelled as an isometry $\mathcal{I} : \mathcal{F}_f \otimes \mathcal{H}_C \to \mathcal{H}_S \otimes \mathcal{H}_C$, given by

$$
\mathcal{I}\,|a', b'\rangle_f\,|c_{a,b}\rangle_C = |0\rangle_S\,|e^{a,b}_{a',b'}\rangle_C + |1\rangle_S\,|f^{a,b}_{a',b'}\rangle_C + |v\rangle_S\,|g^{a,b}_{a',b'}\rangle_C + |m\rangle_S\,|h^{a,b}_{a',b'}\rangle_C ,
\tag{13}
$$

where states $|e^{a,b}_{a',b'}\rangle_C, |f^{a,b}_{a',b'}\rangle_C, |g^{a,b}_{a',b'}\rangle_C, |h^{a,b}_{a',b'}\rangle_C \in \mathcal{H}_C$ are again not necessarily normalized, nor orthogonal. Note that, due to the action of U_1, the photon numbers a, b are no longer correlated to $a', b' \in \{0, 1, 2\}$; nevertheless, we still have $a' + b' \leq 2$. From this, one obtains the final state between the users and the server, $|\phi_2\rangle_{ABSC} = \mathcal{I}\,|\phi_1\rangle_{ABfC}$. Using Theorem 2, one can lower bound the conditional entropy $S(A|C)$, as explained in detail in the next section.

B Two particular attacks

B.1 Adaptive attack with a single photon

The adaptive attack with a single photon that is fed in an agent's laboratory several times during a single round of the key distribution protocol is based on the interaction-free measurement proposed in [54], depict in Figure 5. An agent, say Alice, is placed in one arm of an interferometer which consist of an input polarizing beam splitter and standard balanced beam splitter on its output. Before entering the interferometer, the initial polarization state, say horizontal state $|\psi_0\rangle = |H\rangle$, is rotated by a certain angle θ, so that before the polarizing beam splitter it is $|\psi_\theta\rangle = \cos\theta\,|H\rangle + \sin\theta\,|V\rangle$. In case Alice decided to "reflect", at the output of the interferometer the polarization state of the photon will stay the same, $|\psi_\theta\rangle$. In case she decided to "detect", with probability $\sin^2\theta$ the photon will end up in Alice's laboratory and be absorbed, while with probability $\cos^2\theta$ it will leave the interferometer in polarization state $|\psi_0\rangle$. In the case of the latter, the process is repeated, up to M times. If the rotation angle is chosen to be $\theta = \pi/2M$, after M iterations the polarization state will be $|\psi_{\pi/2}\rangle = |V\rangle$ in case Alice decided to "reflect", while it will stay "frozen" to $|\psi_0\rangle = |H\rangle$ in case she decided to "detect", i.e., the two states will be fully distinguishable, and Eve would know Alice's action. The probability that a photon will not end in Alice's arm M consecutive times when she decided to "detect" is $p = \cos^{2M}\theta = (\cos\frac{\pi}{2M})^{2M}$, which for large M behaves like $p \sim 1 - \pi^2/4M \to 1$. Thus, with probability arbitrarily close to 1 Eve can learn Alice's action without triggering her detector ("activating the bomb" from the original scenario discussed in [54]).

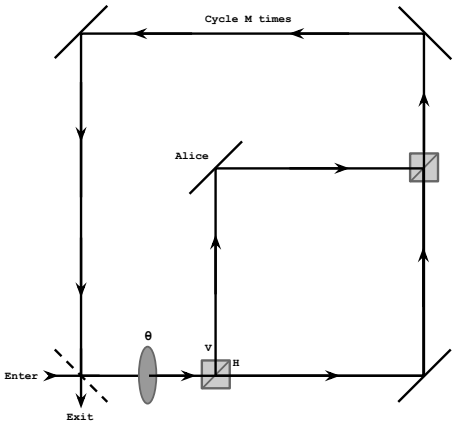


Figure 5: **Eve's attack based on interaction-free measurement.**

But in our case, the aim of Eve is to simulate, as much as possible, the honest scenario, in which Alice's detector will click in about half of the cases. This is achieved with pretty good accuracy for $M = 4$ already, as we have that $p = (\cos\frac{\pi}{8})^8 \approx 0.53$. Nevertheless, if Eve wanted to learn the actions of both agents, she would need to perform two such measurements performed on both agents. But this would inevitably lead to increased double clicks in rounds when both agents decide to "detect" (Note that in order to learn the action of a single agent, say Alice, Eve should perform measurement after her laboratory, thus destroying any possible coherence between photon(s) state in Alice's and Bob's labs). During the parameter estimation phase, Alice and Bob can infer such increased probability of coincidences, and thus detect eavesdropping.

B.2 Multi-photon attack

This is a version of the above interaction attack in which instead of sending a single photon through the interferometer M times, Eve sends M photons only once, in order to learn action of a single agent. Thus, it suffers from the same deficiency as the previous attack: Alice's photon detection is not correlated with Bob's one and therefore will change the joint detection statistics. Again, note that in order to learn the action of a single agent, Eve must perform her measurement on the photons outside her/his lab, thus destroying any possible coherence. In other words, sending a coherent superposition between photon states sent to Alice and Bob offers no advantage.

But this attack features additional problem, in that Eve cannot fully distinguish between an agent's actions, leading her to announce inconsistent messages allowing Alice and Bob to additionally detect cheating. Let us first describe this attack in more detail. Eve sends a multi-photon state $|\Psi_\theta(M)\rangle = |\psi_\theta\rangle^{\otimes M} = (\cos\theta\,|H\rangle + \sin\theta\,|V\rangle)^{\otimes M}$. If Alice decides to "reflect", Eve will receive the same M-photon state $|\Psi_\theta(M)\rangle$ at the output of the interferometer. In case she decides to "detect" and at least one of the photons ends in her arm, there will be less then M photons at the output of the interferometer, and Eve can thus infer Alice's action. But if not a single photon gets detected by Alice, at the output of the interferometer we would have the M-photon state $|\Psi_0(M)\rangle = |\psi_0\rangle^{\otimes M} = |H\rangle^{\otimes M}$. Thus, Eve cannot distinguish the two actions by measuring the photon number, and she needs to subsequently perform polarization measurement. The optimal discrimination probability for the two states is given in terms of the transition probability $\tilde{p} = |\langle\Psi_0|\Psi_\theta\rangle|^2 = |\langle\psi_0|\psi_\theta\rangle|^{2M} = \cos^{2M}\theta$, which is precisely the probability that in the case of deciding to "detect" none of M photons end up in Alice's arm. On the other hand, as before we want that this probability is equal to $1/2$, to match the honest scenario. Thus, if Eve wants to emulate the honest scenario, she must set θ such that the output polarization states are far from fully distinguishable. In other words, the adversary will necessarily occasionally announce messages that are inconsistent with the agents' actions, thus revealing eavesdropping. One can straightforwardly apply our methodology to this case to obtain quantitative expression for the secret key rate. Therefore, we omit this rather complex, but straightforward analysis.

C Security Analysis - Ideal case

In Appendix D, we show how to prove the security of our protocol in the general case, assuming practical devices. To develop the intuition behind the proof in that section, however, we first consider the ideal case scenario. Here, we assume that the server has a perfect single-photon source, Alice's and Bob's detectors are

perfect, which means they have 100% detection efficiency and zero dark counts, but there may be channel loss. Therefore, the perfect single photon state that Alice and Bob expect to be sent is

$$|\phi_0\rangle_f = \left(\frac{\hat{a}^\dagger + \hat{b}^\dagger}{\sqrt{2}}\right)|0,0\rangle_f = \frac{|1,0\rangle_f + |0,1\rangle_f}{\sqrt{2}}, \tag{14}$$

with $|1,0\rangle_f$ and $|0,1\rangle_f$ representing the photon located in Alice's and Bob's arms, respectively. However, we assume that the following entangled state is sent to Alice and Bob by the server (or Eve)

$$|\phi_0\rangle_{fC} = |0,0\rangle_f \otimes |c_{0,0}\rangle + |1,0\rangle_f \otimes |c_{1,0}\rangle_C + |0,1\rangle_f \otimes |c_{0,1}\rangle_C \tag{15}$$

where $|c_{a,b}\rangle_C \in \mathcal{H}_C$ are not necessarily orthogonal nor normalized. Note that, this is the state arriving at A and B's lab, and so it also incorporates channel loss in the $|0,0\rangle_f \otimes |c_{0,0}\rangle$ term. Moreover, as per usual in QKD security proofs, Alice and Bob can enforce symmetry, and so, we may assume $\langle c_{0,1}|c_{0,1}\rangle_C = \langle c_{1,0}|c_{1,0}\rangle$. Therefore, we can write the joint initial state as

$$\begin{aligned} |\phi_0\rangle_{ABfC} &= |\phi_0\rangle_{AB} \otimes |\phi_0\rangle_{fC} \\ &= \frac{1}{2}\Big(|D_v,R\rangle_{AB} + |R,D_v\rangle_{AB} + |D_v,D_v\rangle_{AB} + |R,R\rangle_{AB}\Big) \\ &\quad \otimes \Big(|0,0\rangle_f |c_{0,0}\rangle_C + |1,0\rangle_f |c_{1,0}\rangle_C + |0,1\rangle_f |c_{0,1}\rangle_C\Big). \end{aligned} \tag{16}$$

Alice's and Bob's actions on a given initial photon state are given by

$$\begin{aligned} &|D_v,R\rangle|1,0\rangle \to |D_c,R\rangle|0,0\rangle, && |R,D_v\rangle|1,0\rangle \to |R,D_v\rangle|1,0\rangle, \\ &|D_v,R\rangle|0,1\rangle \to |D_v,R\rangle|0,1\rangle, && |R,D_v\rangle|0,1\rangle \to |R,D_c\rangle|0,0\rangle, \\ &|D_v,D_v\rangle|1,0\rangle \to |D_c,D_v\rangle|0,0\rangle, && |R,R\rangle|1,0\rangle \to |R,R\rangle|1,0\rangle, \\ &|D_v,D_v\rangle|0,1\rangle \to |D_v,D_c\rangle|0,0\rangle, && |R,R\rangle|0,1\rangle \to |R,R\rangle|0,1\rangle, \\ &|D_v,D_v\rangle|0,0\rangle \to |D_v,D_v\rangle|0,0\rangle, && |R,R\rangle|0,0\rangle \to |R,R\rangle|0,0\rangle, \\ &|D_v,R\rangle|0,0\rangle \to |D_v,R\rangle|0,0\rangle, && |R,D_v\rangle|0,0\rangle \to |R,D_v\rangle|0,0\rangle, \end{aligned} \tag{17}$$

and, therefore

$$\begin{aligned} |\phi_1\rangle_{ABfC} = \frac{1}{2}\Big[&|D_c,R\rangle|0,0\rangle|c_{1,0}\rangle + |D_v,R\rangle|0,1\rangle|c_{0,1}\rangle + |R,D_v\rangle|1,0\rangle|c_{1,0}\rangle_C + |R,D_c\rangle|0,0\rangle|c_{0,1}\rangle \\ &+ |D_c,D_v\rangle|0,0\rangle|c_{1,0}\rangle + |D_v,D_c\rangle|0,0\rangle|c_{0,1}\rangle + |R,R\rangle\big(|1,0\rangle|c_{1,0}\rangle + |0,1\rangle|c_{0,1}\rangle \\ &+ |D_v,D_v\rangle|0,0\rangle|c_{0,0}\rangle + |D_v,R\rangle|0,0\rangle|c_{0,0}\rangle + |R,D_v\rangle|0,0\rangle|c_{0,0}\rangle + |R,R\rangle|0,0\rangle|c_{0,0}\rangle\Big]. \end{aligned} \tag{18}$$

Following this, as in the experimental case, the adversary will apply a quantum instrument to the returning photon state which, as before, can be modeled as an isometry, whose action is defined as

$$\mathcal{I}|a',b'\rangle_f |c_{a,b}\rangle_C = |0\rangle_S |e^{a,b}_{a',b'}\rangle_C + |1\rangle_S |f^{a,b}_{a',b'}\rangle_C + |v\rangle_S |g^{a,b}_{a',b'}\rangle_C, \tag{19}$$

where states from $\mathcal{H}_C$ are not necessarily normalized nor orthogonal, and a, b are no longer correlated with a', b' due to Alice's and Bob's actions given by Equation (17). Note that since we are assuming an ideal case, the term corresponding to the message "m" is absent from the above equation.

We are interested only in the rounds when the server announces "1" and neither Alice nor Bob detect a photon, and the users generate the key. Thus, while writing the state after the server applies $\mathcal{I}$ on $|\phi_1\rangle_{ABfC}$, we will omit writing the server's message state $|1\rangle_S$ (corresponding to announcing a result "1"). The final density operator representing the state of the system ABC, conditioned on the event that the server sends the message "1" and none of the users detects a photon (only the rounds used for key generation), is

$$|\phi_2\rangle_{ABC} = \frac{1}{\sqrt{\mathcal{N}}}\Big\{|D_v,R\rangle \otimes |k_{0,0}\rangle + |R,D_v\rangle \otimes |k_{1,1}\rangle + |R,R\rangle \otimes |k_{1,0}\rangle + |D_v,D_v\rangle \otimes |k_{0,1}\rangle\Big\}, \tag{20}$$

where the states $|k_{i,j}\rangle_C$ are associated to Alice establishing the value i and Bob j as a key bit, are given by

$$\begin{aligned}
|k_{0,0}\rangle_C &= \frac{1}{2}\left[|f_{0,1}^{0,1}\rangle + |f_{0,0}^{0,0}\rangle\right], \\
|k_{1,1}\rangle_C &= \frac{1}{2}\left[|f_{1,0}^{1,0}\rangle + |f_{0,0}^{0,0}\rangle\right], \\
|k_{0,1}\rangle_C &= \frac{1}{2}|f_{0,0}^{0,0}\rangle, \\
|k_{1,0}\rangle_C &= \frac{1}{2}\left[|f_{1,0}^{1,0}\rangle + |f_{0,1}^{0,1}\rangle + |f_{0,0}^{0,0}\rangle\right].
\end{aligned} \tag{21}$$

Note that, though we are assuming in this ideal setting, that A and B's devices are ideal, the adversarial server may still "simulate" imperfect detectors which may have, for instance, dark counts (incorporated in the term $\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle$ which is the probability the server sends a positive message in the event a vacuum actually enters its lab). The normalization constant $\mathcal{N}$ is, again, the probability to obtain the result 1, p(1), when there were no clicks at the users' detectors, and is given by

$$\mathcal{N} = \langle k_{0,0}|k_{0,0}\rangle + \langle k_{1,1}|k_{1,1}\rangle + \langle k_{1,0}|k_{1,0}\rangle + \langle k_{0,1}|k_{0,1}\rangle = \mathrm{p}(1). \tag{22}$$

As in the experimental case, we define $\mathrm{p}_{0,0} = \mathrm{p}(D_v,R\,;1) = \langle k_{0,0}|k_{0,0}\rangle$ as the joint probability for the event when Alice detects vacuum and Bob reflects, and the server announces the result "1", and analogously $\mathrm{p}_{1,1}$, $\mathrm{p}_{0,1}$ and $\mathrm{p}_{1,0}$. Again, we use the semicolon (;) to denote logical AND operation between two propositions. Therefore, we can define the probability to share the key as $\mathrm{p}_{key} = \mathrm{p}_{0,0} + \mathrm{p}_{1,1}$ and the probability of an error as $\mathrm{p}_{err} = \mathrm{p}_{0,1} + \mathrm{p}_{1,0}$. When we evaluate our key rate bound, we use $\mathcal{Q}$ to be the probability that the server announces the result "1", given both Alice and Bob reflected, conditioned on a photon arriving at the server. Finally, we allow the adversarial server to "simulate" dark counts at a rate of p_d (to its advantage), and we use T to mean the probability of transmittance in one direction, namely $1-T$ is the probability the photon is dropped before it gets to A or B (the probability the photon returns to the server if A and B reflect is T^2). In the ideal case, it is easy to see that

$$\begin{aligned}
\mathrm{p}_{0,0} = \langle k_{0,0}|k_{0,0}\rangle &= \frac{1}{4}\left(\frac{T^2}{4} + \frac{T(1-T)p_d}{2}\right), \quad \mathrm{p}_{0,1} = \langle k_{0,1}|k_{0,1}\rangle = \frac{(1-T)p_d}{4}, \\
\mathrm{p}_{1,1} = \langle k_{1,1}|k_{1,1}\rangle &= \frac{1}{4}\left(\frac{T^2}{4} + \frac{T(1-T)p_d}{2}\right), \quad \mathrm{p}_{1,0} = \langle k_{1,0}|k_{1,0}\rangle = \frac{\mathcal{Q}\cdot T^2}{4}.
\end{aligned} \tag{23}$$

Expanding $\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle$, needed for the entropy bound computation, we find:

$$\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle = \frac{1}{4}(\mathrm{Re}\,\langle f_{0,1}^{0,1}|f_{1,0}^{1,0}\rangle + \mathrm{Re}\,\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle + \mathrm{Re}\,\langle f_{0,1}^{0,1}|f_{0,0}^{0,0}\rangle + \mathrm{Re}\,\langle f_{0,0}^{0,0}|f_{1,0}^{1,0}\rangle) \tag{24}$$

Expanding $\langle k_{1,0}|k_{1,0}\rangle$ we find:

$$\begin{aligned}
\langle k_{1,0}|k_{1,0}\rangle &= \frac{1}{4}(\langle f_{1,0}^{1,0}|f_{1,0}^{1,0}\rangle + \langle f_{0,1}^{0,1}|f_{0,1}^{0,1}\rangle + \langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle + 2\mathrm{Re}(\langle f_{0,1}^{0,1}|f_{1,0}^{1,0}\rangle + \langle f_{0,1}^{0,1}|f_{0,0}^{0,0}\rangle + \langle f_{1,0}^{1,0}|f_{0,0}^{0,0}\rangle)). \\
\Rightarrow \mathrm{Re}(\langle f_{0,1}^{0,1}|f_{1,0}^{1,0}\rangle + \langle f_{0,1}^{0,1}|f_{0,0}^{0,0}\rangle + \langle f_{1,0}^{1,0}|f_{0,0}^{0,0}\rangle) &= \frac{4\,\langle k_{1,0}|k_{1,0}\rangle - (\langle f_{1,0}^{1,0}|f_{1,0}^{1,0}\rangle + \langle f_{0,1}^{0,1}|f_{0,1}^{0,1}\rangle + \langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle)}{2}
\end{aligned} \tag{25}$$

Substituting Equation (25) into Equation (24) we have:

$$\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle = \frac{1}{2}\,\langle k_{1,0}|k_{1,0}\rangle - \frac{1}{8}(\langle f_{1,0}^{1,0}|f_{1,0}^{1,0}\rangle + \langle f_{0,1}^{0,1}|f_{0,1}^{0,1}\rangle + \langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle). \tag{26}$$

The term $\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle$ is an observable quantity, it is simply $4\,\langle k_{0,1}|k_{0,1}\rangle = 4p_{0,1}$. The values of $\langle f_{0,1}^{0,1}|f_{0,1}^{0,1}\rangle$ and $\langle f_{1,0}^{1,0}|f_{1,0}^{1,0}\rangle$ can be bounded by solving the following quadratic equation (derived from the expansion of $\langle k_{0,0}|k_{0,0}\rangle$ and $\langle k_{1,1}|k_{1,1}\rangle$ respectively):

$$\langle f_{0,1}^{0,1}|f_{0,1}^{0,1}\rangle + 2\cos\theta\sqrt{\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle}\sqrt{\langle f_{0,1}^{0,1}|f_{0,1}^{0,1}\rangle} + (\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle - 4\,\langle k_{0,0}|k_{0,0}\rangle). \tag{27}$$

Similarly for $\langle f_{1,0}^{1,0}|f_{1,0}^{1,0}\rangle$. This allows us to minimize $|\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle|$, thus minimizing the adversary's uncertainty (i.e., minimizing $S(A|C)$).

At this point, we compute the conditional entropy between Alice and the adversary, $S(A|C)$, for the rounds where raw key bits are generated. Using Equation (20), the density operator, after dropping off-diagonal terms, with $|k_{i,j}\rangle_C\langle k_{l,m}|$, for $(i,j) \neq (l,m)$, is

$$\rho_{ABC} = \frac{1}{\mathcal{N}} \Big(|D_v,R\rangle_{AB}\langle D_v,R| \otimes |k_{0,0}\rangle_C\langle k_{0,0}| + |R,D_v\rangle_{AB}\langle R,D_v| \otimes |k_{1,1}\rangle_C\langle k_{1,1}| + |R,R\rangle_{AB}\langle R,R| \otimes |k_{1,0}\rangle_C\langle k_{1,0}| + |D_v,D_v\rangle_{AB}\langle D_v,D_v| \otimes |k_{0,1}\rangle_C\langle k_{0,1}| \Big). \tag{28}$$

The state $|D_v,R\rangle\langle D_v,R|$, describing Alice detecting without a click and Bob reflecting, is associated to a shared key bit 0. Similarly, $|R,D_v\rangle\langle R,D_v|$ is associated to a key bit 1. Whereas, $|R,R\rangle\langle R,R|$ and $|D_v,D_v\rangle\langle D_v,D_v|$ corresponds to errors in the key, when the two users establish opposite key bit values.

Now that we have a description of the quantum state, we can use Theorem 2 to compute a bound on the conditional entropy $S(A|C)$ leading us to:

$$S(A|C) \geq \frac{\langle k_{0,0}|k_{0,0}\rangle + \langle k_{1,1}|k_{1,1}\rangle}{\mathcal{N}} \left[h\left(\frac{\langle k_{0,0}|k_{0,0}\rangle}{\langle k_{0,0}|k_{0,0}\rangle + \langle k_{1,1}|k_{1,1}\rangle} \right) - h(\lambda_0) \right], \tag{29}$$

with λ_0 is defined as in Equation (37).

We present the dependence of the secret key rate r on the total number of rounds N for different values of $\mathcal{Q}$ (including the one obtained from the experimental set-up) in Figure 6 for $T = 1$. Other parameters are taken from [31] as $\epsilon = 10^{-5}$, $\epsilon_{EC} = 10^{-10}$ and $\epsilon' = 10^{-7}$. We also assume $\epsilon_{PE} = 10^{-11}$. In Figure 7, we report key-rate as a function of total transmission loss in one direction where we set p_d to be a negligible 10^{-8} to consider ideal devices on the server also.

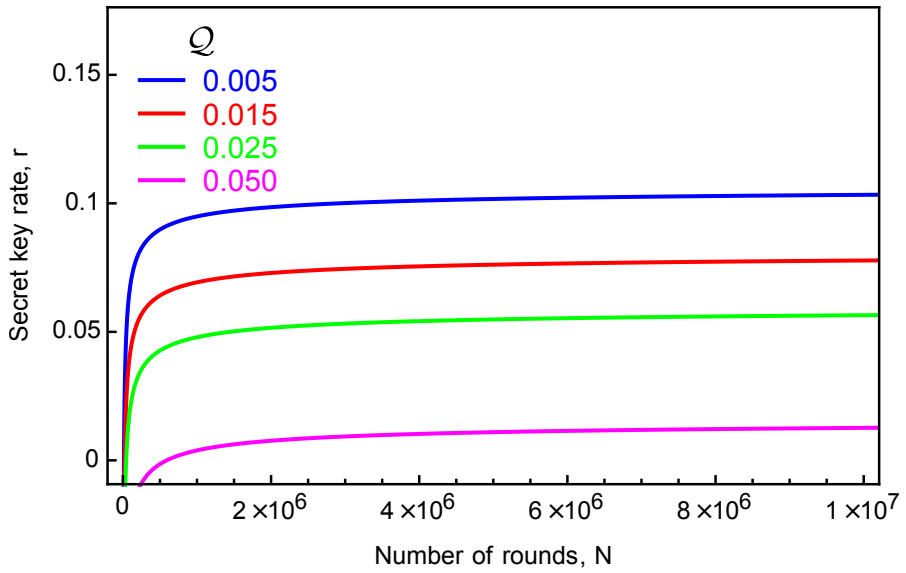


Figure 6: The secret key rate r is plotted against N, for the ideal case of perfect single-photon sources and detectors. The blue, green and magenta curves correspond to the values of $\mathcal{Q}$ to be $0.005, 0.025$ and 0.05, respectively. Whereas, the red curve represents the experimentally observed value of $\mathcal{Q}$, 0.015.

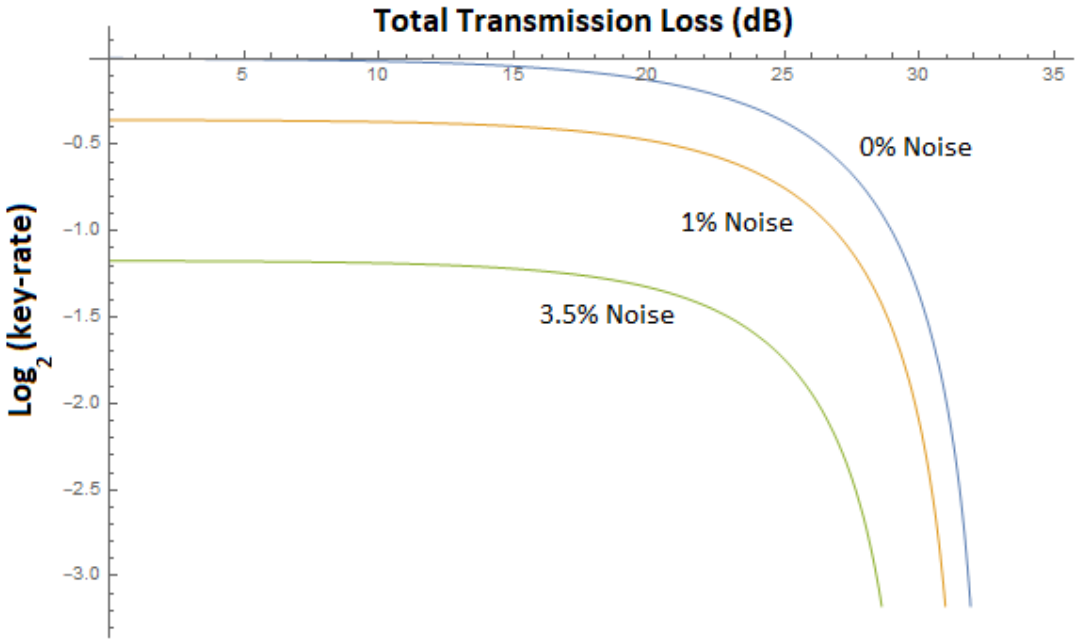


Figure 7: The secret key rate r is plotted against transmission loss in one direction (server to end-users), for the ideal case of perfect single-photon sources and detectors in the asymptotic setting (taking the number of iterations to be infinite and with perfect error correction).

D Security Analysis - General Case

By straightforward algebra, from Equations (11), (12) and (13), we get $|\phi_2\rangle_{ABSC} = \mathcal{I}\,|\phi_1\rangle_{ABfC}$. However, we are only interested in the key-generation rounds, i.e., we condition to the event when the server announces "1" and neither Alice nor Bob receives a click. Hence, omitting writing the message state $|1\rangle_S$, the final density operator (without the off-diagonal terms) of the system ABC is

$$\begin{aligned}\rho_{ABC} = \frac{1}{\mathcal{N}}\Big[& |D_v,R\rangle_{AB}\langle D_v,R| \otimes |k_{0,0}\rangle_C\langle k_{0,0}| + |R,D_v\rangle_{AB}\langle R,D_v| \otimes |k_{1,1}\rangle_C\langle k_{1,1}| \\ & + |D_\ell,R\rangle_{AB}\langle D_\ell,R| \otimes |k^1_{0,0}\rangle_C\langle k^1_{0,0}| + |R,D_\ell\rangle_{AB}\langle R,D_\ell| \otimes |k^1_{1,1}\rangle_C\langle k^1_{1,1}| \\ & + |D'_\ell,R\rangle_{AB}\langle D'_\ell,R| \otimes |k^2_{0,0}\rangle_C\langle k^2_{0,0}| + |R,D'_\ell\rangle_{AB}\langle R,D'_\ell| \otimes |k^2_{1,1}\rangle_C\langle k^2_{1,1}| \\ & + |D_\ell D'_\ell,R\rangle_{AB}\langle D_\ell D'_\ell,R| \otimes |k^3_{0,0}\rangle_C\langle k^3_{0,0}| + |R,D_\ell D'_\ell\rangle_{AB}\langle R,D_\ell D'_\ell| \otimes |k^3_{1,1}\rangle_C\langle k^3_{1,1}| \\ & + |D_v,D_v\rangle_{AB}\langle D_v,D_v| \otimes |k_{0,1}\rangle_C\langle k_{0,1}| + |R,R\rangle_{AB}\langle R,R| \otimes |k_{1,0}\rangle_C\langle k_{1,0}| \\ & + |D_\ell,D_v\rangle_{AB}\langle D_\ell,D_v| \otimes |k^1_{0,1}\rangle_C\langle k^1_{0,1}| + |D_v,D_\ell\rangle_{AB}\langle D_v,D_\ell| \otimes |k^2_{0,1}\rangle_C\langle k^2_{0,1}| \\ & + |D_\ell,D'_\ell\rangle_{AB}\langle D_\ell,D'_\ell| \otimes |k^3_{0,1}\rangle_C\langle k^3_{0,1}| + |D'_\ell,D_\ell\rangle_{AB}\langle D'_\ell,D_\ell| \otimes |k^4_{0,1}\rangle_C\langle k^4_{0,1}| \\ & + |D_\ell D'_\ell,D_v\rangle_{AB}\langle D_\ell D'_\ell,D_v| \otimes |k^5_{0,1}\rangle_C\langle k^5_{0,1}| + |D_v,D_\ell D'_\ell\rangle_{AB}\langle D_v,D_\ell D'_\ell| \otimes |k^6_{0,1}\rangle_C\langle k^6_{0,1}|\Big]. \end{aligned} \tag{30}$$

Note that, as before, we use commas in the states from $\mathcal{H}_A \otimes \mathcal{H}_B$ to separate the quantum numbers defining Alice's and Bob's apparatus states: $|D_\ell D'_\ell,R\rangle_{AB}$ means that Alice opted to detect, unsuccessfully (due to finite detection efficiency) the two photons present in her lab, while Bob set his apparatus to reflect, etc. The states $|k_{i,j}\rangle_C$, etc., are associated to the cases when Alice establishes the value i and Bob j as a key bit, and are given

by

$$
\begin{aligned}
|k_{0,0}\rangle &= \frac{1}{2}\left[|f^{0,0}_{0,0}\rangle + |f^{0,1}_{0,1}\rangle + |f^{0,2}_{0,2}\rangle\right], & |k_{1,1}\rangle &= \frac{1}{2}\left[|f^{0,0}_{0,0}\rangle + |f^{1,0}_{1,0}\rangle + |f^{2,0}_{2,0}\rangle\right],\\
|k^1_{0,0}\rangle &= \frac{1}{2}\sqrt{\mathrm{p}^A_\ell}\left[|f^{1,0}_{0,0}\rangle + |f^{1,1}_{0,1}\rangle\right], & |k^1_{1,1}\rangle &= \frac{1}{2}\sqrt{\mathrm{p}^B_\ell}\left[|f^{0,1}_{0,0}\rangle + |f^{1,1}_{1,0}\rangle\right],\\
|k^2_{0,0}\rangle &= \frac{1}{2}\sqrt{\mathrm{p}^A_\ell}\,|f^{1',1}_{0,1}\rangle, & |k^2_{1,1}\rangle &= \frac{1}{2}\sqrt{\mathrm{p}^B_\ell}\,|f^{1,1'}_{1,0}\rangle,\\
|k^3_{0,0}\rangle &= \frac{1}{2}\mathrm{p}^A_\ell\,|f^{2,0}_{0,0}\rangle, & |k^3_{1,1}\rangle &= \frac{1}{2}\mathrm{p}^B_\ell\,|f^{0,2}_{0,0}\rangle,\\
|k_{0,1}\rangle &= \frac{1}{2}\,|f^{0,0}_{0,0}\rangle, & |k_{1,0}\rangle &= \frac{1}{2}\big[\,|f^{0,0}_{0,0}\rangle + |f^{1,0}_{1,0}\rangle + |f^{0,1}_{0,1}\rangle + |f^{2,0}_{2,0}\rangle,\\
& & & \quad + |f^{1,1'}_{1,1'}\rangle + |f^{1',1}_{1',1}\rangle + |f^{0,2}_{0,2}\rangle\,\big],\\
|k^1_{0,1}\rangle &= \frac{1}{2}\sqrt{\mathrm{p}^A_\ell}\,|f^{1,0}_{0,0}\rangle, & |k^4_{0,1}\rangle &= \frac{1}{2}\sqrt{\mathrm{p}^A_\ell\,\mathrm{p}^B_\ell}\,|f^{1',1}_{0,0}\rangle,\\
|k^2_{0,1}\rangle &= \frac{1}{2}\sqrt{\mathrm{p}^B_\ell}\,|f^{0,1}_{0,0}\rangle, & |k^5_{0,1}\rangle &= \frac{1}{2}\mathrm{p}^A_\ell\,|f^{2,0}_{0,0}\rangle,\\
|k^3_{0,1}\rangle &= \frac{1}{2}\sqrt{\mathrm{p}^A_\ell\,\mathrm{p}^B_\ell}\,|f^{1,1'}_{0,0}\rangle, & |k^6_{0,1}\rangle &= \frac{1}{2}\mathrm{p}^B_\ell\,|f^{0,2}_{0,0}\rangle.
\end{aligned}
\tag{31}
$$

Above, as well as in rest of the Appendix, for simplicity we omit writing the labels of the quantum states (A, B, C, S and f), whenever it is implicitly unambiguous to which space they belong by their quantum numbers (D_v, $0,0$, etc.).

The normalization constant $\mathcal{N}$ from Equation (30) is the probability to obtain the result "1" when there were no clicks at the agents' detectors, given by

$$
\begin{aligned}
\mathcal{N} = \;& \langle k_{0,0}|k_{0,0}\rangle + \langle k^1_{0,0}|k^1_{0,0}\rangle + \langle k^2_{0,0}|k^2_{0,0}\rangle + \langle k^3_{0,0}|k^3_{0,0}\rangle + \langle k_{1,1}|k_{1,1}\rangle + \langle k^1_{1,1}|k^1_{1,1}\rangle + \langle k^2_{1,1}|k^2_{1,1}\rangle + \langle k^3_{1,1}|k^3_{1,1}\rangle\\
& + \langle k_{0,1}|k_{0,1}\rangle + \langle k^1_{0,1}|k^1_{0,1}\rangle + \langle k^2_{0,1}|k^2_{0,1}\rangle + \langle k^3_{0,1}|k^3_{0,1}\rangle + \langle k^4_{0,1}|k^4_{0,1}\rangle + \langle k^5_{0,1}|k^5_{0,1}\rangle + \langle k^6_{0,1}|k^6_{0,1}\rangle + \langle k_{1,0}|k_{1,0}\rangle\,.
\end{aligned}
\tag{32}
$$

In ρ_{ABC}, given by Equation (30), the state $|D_v,R\rangle\langle D_v,R|$ describes Alice detecting without a click and Bob reflecting, and is associated to a shared key bit of 0. Let us define $\mathrm{p}_{0,0} = \mathrm{p}(D_v,R\,;1) = \langle k_{0,0}|k_{0,0}\rangle$ as the joint probability for the event when Alice detects vacuum and Bob reflects, and the server announces the result "1", which corresponds to the users sharing a key bit of 0. Here we use the semicolon (;) to denote logical AND operation between two propositions. Note that $|D_\ell,R\rangle\langle D_\ell,R|$, $|D'_\ell,R\rangle\langle D'_\ell,R|$, and $|D_\ell D'_\ell,R\rangle\langle D_\ell D'_\ell,R|$ also correspond to a shared key bit of 0, and are a consequence of Alice's imperfect detector and multi-photon events. Therefore, one can analogously define the probabilities $\mathrm{p}^1_{0,0}, \mathrm{p}^2_{0,0}$ and $\mathrm{p}^3_{0,0}$, such that the total probability of the users sharing a key bit of 0 can be given by $\tilde{\mathrm{p}}_{0,0} = \mathrm{p}_{0,0} + \mathrm{p}^1_{0,0} + \mathrm{p}^2_{0,0} + \mathrm{p}^3_{0,0}$. Analogously, the probabilities, $\mathrm{p}_{1,1}, \mathrm{p}^1_{1,1}, \mathrm{p}^2_{1,1}$ and $\mathrm{p}^3_{1,1}$, associated to a key bit 1 are defined. The k_{ij}'s with $i \neq j$ are associated to the errors, i.e., when the two users establish opposite key bit values. From the above definitions, using k_{ij} and $\mathcal{N}$, we have

$$
\frac{\langle k_{0,0}|k_{0,0}\rangle + \langle k^1_{0,0}|k^1_{0,0}\rangle + \langle k^2_{0,0}|k^2_{0,0}\rangle + \langle k^3_{0,0}|k^3_{0,0}\rangle}{\mathcal{N}} = \mathrm{p}(D_v,R \vee D_\ell,R \vee D'_\ell,R \vee D_\ell D'_\ell,R|1). \tag{33}
$$

Here, by $\mathrm{p}(\mathcal{P}|\mathcal{C})$ we denote the conditional probability that the proposition $\mathcal{P}$ holds (in the above case, Alice detects and observes no clicks, while Bob reflects), given that the condition $\mathcal{C}$ is satisfied (in the above case, the server announces "1"). Therefore, using the following terminology for different probabilities (to be used in parameter estimation described in the next section), the probability to share the key is given by

$$
\begin{aligned}
\mathrm{p}_{key} &= \left[\langle k_{0,0}|k_{0,0}\rangle + \langle k^1_{0,0}|k^1_{0,0}\rangle + \langle k^2_{0,0}|k^2_{0,0}\rangle + \langle k^3_{0,0}|k^3_{0,0}\rangle\right] + \left[\langle k_{1,1}|k_{1,1}\rangle + \langle k^1_{1,1}|k^1_{1,1}\rangle + \langle k^2_{1,1}|k^2_{1,1}\rangle + \langle k^3_{1,1}|k^3_{1,1}\rangle\right]\\
&= \left[\mathrm{p}_{0,0} + \mathrm{p}^1_{0,0} + \mathrm{p}^2_{0,0} + \mathrm{p}^3_{0,0}\right] + \left[\mathrm{p}_{1,1} + \mathrm{p}^1_{1,1} + \mathrm{p}^2_{1,1} + \mathrm{p}^3_{1,1}\right]\\
&= \tilde{\mathrm{p}}_{0,0} + \tilde{\mathrm{p}}_{1,1}\\
&= \mathrm{p}(D_v,R \vee D_\ell,R \vee D'_\ell,R \vee D_\ell D'_\ell,R\,;1) + \mathrm{p}(R,D_v \vee R,D_\ell \vee R,D'_\ell \vee R,D_\ell D'_\ell\,;1),
\end{aligned}
\tag{34}
$$

where $\mathrm{p}(D_v,R \vee D_\ell,R \vee D'_\ell,R \vee D_\ell D'_\ell,R\,;1)$ represents the joint probability of the following event: Alice detects vacuum, Bob reflects, and the server announces the result "1"; and analogously for the other term. As before, we use the semicolon (;) to denote logical AND operation between two propositions, instead of introducing the

additional parenthesis for the first one, and using the standard symbol $\wedge$. The probability of error in the raw key is given by

$$\begin{aligned}
\mathrm{p}_{err} &= \left[\langle k_{0,1}|k_{0,1}\rangle + \langle k^1_{0,1}|k^1_{0,1}\rangle + \langle k^2_{0,1}|k^2_{0,1}\rangle + \langle k^3_{0,1}|k^3_{0,1}\rangle + \langle k^4_{0,1}|k^4_{0,1}\rangle + \langle k^5_{0,1}|k^5_{0,1}\rangle + \langle k^6_{0,1}|k^6_{0,1}\rangle\right] + \langle k_{1,0}|k_{1,0}\rangle \\
&= \left[\mathrm{p}_{0,1} + \mathrm{p}^1_{0,1} + \mathrm{p}^2_{0,1} + \mathrm{p}^3_{0,1} + \mathrm{p}^4_{0,1} + \mathrm{p}^5_{0,1} + \mathrm{p}^6_{0,1}\right] + \mathrm{p}_{1,0} \\
&= \tilde{\mathrm{p}}_{0,1} + \tilde{\mathrm{p}}_{1,0} \\
&= \mathrm{p}(D_v, D_v \vee D_\ell, D_v \vee D_v, D_\ell \vee D_\ell, D'_\ell \vee D'_\ell, D_\ell \vee D_v, D_\ell D'_\ell \vee D'_\ell D_\ell, D_v\,; 1) + \mathrm{p}(RR\,; 1),
\end{aligned} \tag{35}$$

where $\mathrm{p}(D_v, D_v \vee D_\ell, D_v \vee D_v, D_\ell \vee D_\ell, D'_\ell \vee D'_\ell, D_\ell \vee D_v, D_\ell D'_\ell \vee D'_\ell D_\ell, D_v\,; 1)$ represents the joint probability of the event: Alice and Bob both detect vacuum, and that the server announces the result "1"; and analogously for the other term. Note that the probabilities $\tilde{\mathrm{p}}_{i,j}$ can be observed from the experiment directly.

To obtain the secret key rate, we again use the bound given in Theorem 2, as

$$\begin{aligned}
S(A|C) \;\geq\; & \frac{\langle k_{0,0}|k_{0,0}\rangle + \langle k_{1,1}|k_{1,1}\rangle}{\mathcal{N}} \left(h\left[\frac{\langle k_{0,0}|k_{0,0}\rangle}{\langle k_{0,0}|k_{0,0}\rangle + \langle k_{1,1}|k_{1,1}\rangle}\right] - h(\lambda_0)\right) \\
& + \frac{\langle k^1_{0,0}|k^1_{0,0}\rangle + \langle k^1_{1,1}|k^1_{1,1}\rangle}{\mathcal{N}} \left(h\left[\frac{\langle k^1_{0,0}|k^1_{0,0}\rangle}{\langle k^1_{0,0}|k^1_{0,0}\rangle + \langle k^1_{1,1}|k^1_{1,1}\rangle}\right] - h(\lambda_1)\right) \\
& + \frac{\langle k^2_{0,0}|k^2_{0,0}\rangle + \langle k^2_{1,1}|k^2_{1,1}\rangle}{\mathcal{N}} \left(h\left[\frac{\langle k^2_{0,0}|k^2_{0,0}\rangle}{\langle k^2_{0,0}|k^2_{0,0}\rangle + \langle k^2_{1,1}|k^2_{1,1}\rangle}\right] - h(\lambda_2)\right) \\
& + \frac{\langle k^3_{0,0}|k^3_{0,0}\rangle + \langle k^3_{1,1}|k^3_{1,1}\rangle}{\mathcal{N}} \left(h\left[\frac{\langle k^3_{0,0}|k^3_{0,0}\rangle}{\langle k^3_{0,0}|k^3_{0,0}\rangle + \langle k^3_{1,1}|k^3_{1,1}\rangle}\right] - h(\lambda_3)\right) \\
& + \frac{\langle k_{0,1}|k_{0,1}\rangle + \langle k_{1,0}|k_{1,0}\rangle}{\mathcal{N}} \left(h\left[\frac{\langle k_{0,1}|k_{0,1}\rangle}{\langle k_{0,1}|k_{0,1}\rangle + \langle k_{1,0}|k_{1,0}\rangle}\right] - h(\lambda_4)\right),
\end{aligned} \tag{36}$$

where $h(\cdot)$ is the binary Shannon entropy, and λ_i's are defined in the following way

$$\begin{aligned}
\lambda_0 &= \frac{1}{2}\left(1 + \frac{\sqrt{\left(\langle k_{0,0}|k_{0,0}\rangle - \langle k_{1,1}|k_{1,1}\rangle\right)^2 + 4\mathrm{Re}^2\,\langle k_{0,0}|k_{1,1}\rangle}}{\langle k_{0,0}|k_{0,0}\rangle + \langle k_{1,1}|k_{1,1}\rangle}\right), \\
\lambda_1 &= \frac{1}{2}\left(1 + \frac{\sqrt{\left(\langle k^1_{0,0}|k^1_{0,0}\rangle - \langle k^1_{1,1}|k^1_{1,1}\rangle\right)^2 + 4\mathrm{Re}^2\,\langle k^1_{0,0}|k^1_{1,1}\rangle}}{\langle k^1_{0,0}|k^1_{0,0}\rangle + \langle k^1_{1,1}|k^1_{1,1}\rangle}\right), \\
\lambda_2 &= \frac{1}{2}\left(1 + \frac{\sqrt{\left(\langle k^2_{0,0}|k^2_{0,0}\rangle - \langle k^2_{1,1}|k^2_{1,1}\rangle\right)^2 + 4\mathrm{Re}^2\,\langle k^2_{0,0}|k^2_{1,1}\rangle}}{\langle k^2_{0,0}|k^2_{0,0}\rangle + \langle k^2_{1,1}|k^2_{1,1}\rangle}\right), \\
\lambda_3 &= \frac{1}{2}\left(1 + \frac{\sqrt{\left(\langle k^3_{0,0}|k^3_{0,0}\rangle - \langle k^3_{1,1}|k^3_{1,1}\rangle\right)^2 + 4\mathrm{Re}^2\,\langle k^3_{0,0}|k^3_{1,1}\rangle}}{\langle k^3_{0,0}|k^3_{0,0}\rangle + \langle k^3_{1,1}|k^3_{1,1}\rangle}\right), \\
\lambda_4 &= \frac{1}{2}\left(1 + \frac{\sqrt{\left(\langle k_{0,1}|k_{0,1}\rangle - \langle k_{1,0}|k_{1,0}\rangle\right)^2 + 4\mathrm{Re}^2\,\langle k_{0,1}|k_{1,0}\rangle}}{\langle k_{0,1}|k_{0,1}\rangle + \langle k_{1,0}|k_{1,0}\rangle}\right).
\end{aligned} \tag{37}$$

The first four terms in $S(A|C)$ correspond to the keys shared between Alice and Bob, while the last term corresponds to errors in the key. However, we estimate the lower bound on $S(A|C)$ by considering only the first term since its contribution to the entropy is far larger than that of any of the other terms. From the expression (37) for λ_0, we see that minimizing $S(A|C)$ essentially means minimizing $\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle$. Therefore, in addition to different probabilities obtained from the experiment, we also need to estimate $\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle$. We proceed by computing the lower bound for $\mathrm{Re}^2\,\langle k_{0,0}|k_{1,1}\rangle$, i.e., for $|\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle\,|$. Notice that the lower it is, the closer to $1/2$ λ_0 is, i.e., the closer to 1 the $h(\lambda_0)$ is, and the worst case scenario for $S(A|C)$, has the lowest value.

Let us use the following notation for simplification,

$$|x\rangle = |f^{1,0}_{1,0}\rangle + |f^{2,0}_{2,0}\rangle\,, \qquad |y\rangle = |f^{0,1}_{0,1}\rangle + |f^{0,2}_{0,2}\rangle\,, \qquad |z\rangle = |f^{1,1'}_{1,1'}\rangle + |f^{1',1}_{1',1}\rangle\,. \tag{38}$$

We can rewrite $|k_{0,0}\rangle$ and $|k_{1,1}\rangle$ from Equation (31), to obtain $\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle$ as

$$\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle = \frac{1}{4}\left[\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle + \mathrm{Re}\,\langle x|f_{0,0}^{0,0}\rangle + \mathrm{Re}\,\langle f_{0,0}^{0,0}|y\rangle + \mathrm{Re}\,\langle x|y\rangle\right]. \tag{39}$$

From the error term, we have $\langle k_{1,0}|k_{1,0}\rangle = \mathrm{p}(R,R|1)\mathrm{p}(R,R) = \mathcal{Q}/4$, where $\mathcal{Q} = \mathrm{p}(R,R|1)$ is the probability that the server announces the result "1", given both Alice and Bob reflected, and $\mathrm{p}(R,R) = 1/4$. With straightforward substitution from the above into Equation (39), with $\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle = 4\,\langle k_{0,1}|k_{0,1}\rangle = 4\mathrm{p}_{0,1}$, we get

$$\langle k_{0,0}|k_{1,1}\rangle = \frac{\mathcal{Q}}{8} + \frac{\mathrm{p}_{0,1}}{2} - \frac{1}{8}\left[\langle x|x\rangle + \langle y|y\rangle + \langle z|z\rangle\right] - \frac{1}{4}\left[\langle x|z\rangle + \langle y|z\rangle + \langle f_{0,0}^{0,0}|z\rangle\right]. \tag{40}$$

In the ideal case, with no vacuum or multi-photon pulses, when $\langle x|x\rangle = 4\,\langle k_{0,0}|k_{0,0}\rangle = 4\mathrm{p}_{0,0}$ and $\langle y|y\rangle = 4\,\langle k_{1,1}|k_{1,1}\rangle = 4\mathrm{p}_{1,1}$, we recover the ideal case expression 26 from Appendix C. By writing $\langle x|z\rangle = |\,\langle x|z\rangle\,|e^{\varphi_{x,z}}$, we have

$$\mathrm{Re}\,\langle x|z\rangle = |\,\langle x|z\rangle\,|\cos\varphi_{x,y} = ||\,|x\rangle\,|| \cdot ||\,|z\rangle\,|| \cdot |\cos\chi_{x,z}|\cos\varphi_{x,z} = \sqrt{\langle x|x\rangle}\sqrt{\langle z|z\rangle}\cos\theta_{x,z}, \tag{41}$$

where $\chi_{x,z}$ denotes the angle between $|x\rangle$ and $|z\rangle$ and $\cos\theta_{x,z} \equiv |\cos\chi_{x,z}|\cos\varphi_{x,z}$, and analogously for $\mathrm{Re}\,\langle y|z\rangle$ and so on. Therefore, the final expression for $\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle$ is

$$\begin{aligned}\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle = \frac{\mathcal{Q}}{8} + \frac{\mathrm{p}_{0,1}}{2} - \frac{1}{8}\left[\langle x|x\rangle + \langle y|y\rangle + \langle z|z\rangle\right] - \frac{1}{4}\left[\sqrt{\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle}\sqrt{\langle z|z\rangle}\cos\theta_{f,z}\right] \\ -\frac{1}{4}\left[\sqrt{\langle x|x\rangle}\sqrt{\langle z|z\rangle}\cos\theta_{x,z} + \sqrt{\langle y|y\rangle}\sqrt{\langle z|z\rangle}\cos\theta_{y,z}\right].\end{aligned} \tag{42}$$

To obtain $\langle x|x\rangle$ and $\langle y|y\rangle$, consider again $|k_{0,0}\rangle$ and $|k_{1,1}\rangle$ from Equation (31)

$$\begin{aligned}\langle k_{1,1}|k_{1,1}\rangle &= \frac{1}{4}\left[\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle + \langle x|x\rangle + 2\mathrm{Re}\,\langle f_{0,0}^{0,0}|x\rangle\right], \\ \langle k_{0,0}|k_{0,0}\rangle &= \frac{1}{4}\left[\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle + \langle y|y\rangle + 2\mathrm{Re}\,\langle f_{0,0}^{0,0}|y\rangle\right].\end{aligned} \tag{43}$$

Note that $\langle f_{0,0}^{0,0}|f_{0,0}^{0,0}\rangle = 4\,\langle k_{0,1}|k_{0,1}\rangle = 4\mathrm{p}_{0,1}, \langle k_{0,0}|k_{0,0}\rangle = \mathrm{p}_{0,0}$ and $\langle k_{1,1}|k_{1,1}\rangle = \mathrm{p}_{1,1}$. Therefore, solving the quadratic equations obtained from (43), we get the following positive roots of $\sqrt{\langle x|x\rangle}$ and $\sqrt{\langle y|y\rangle}$,

$$\begin{aligned}\sqrt{\langle x|x\rangle} &= 2\left[-\sqrt{\mathrm{p}_{0,1}}\cos\theta_{x,f} + \sqrt{\mathrm{p}_{1,1} - (1 - \cos^2\theta_{x,f})\,\mathrm{p}_{0,1}}\right], \\ \sqrt{\langle y|y\rangle} &= 2\left[-\sqrt{\mathrm{p}_{0,1}}\cos\theta_{y,f} + \sqrt{\mathrm{p}_{0,0} - (1 - \cos^2\theta_{y,f})\,\mathrm{p}_{0,1}}\right].\end{aligned} \tag{44}$$

Analogously, for $\langle z|z\rangle$ we have

$$\begin{aligned}\langle z|z\rangle + \underbrace{2\left[\sqrt{\langle x|x\rangle}\,\cos\,\theta_{x,z} + \sqrt{\langle y|y\rangle}\,\cos\,\theta_{y,z} + 2\sqrt{\mathrm{p}_{0,1}}\,\cos\,\theta_{f,z}\right]}_{\beta}\sqrt{\langle z|z\rangle} \\ \underbrace{+\,4\left[\mathrm{p}_{0,1} - \mathrm{p}_{1,0}\right] + \left[\langle x|x\rangle + \langle y|y\rangle + 2\sqrt{\langle x|x\rangle}\sqrt{\langle y|y\rangle}\,\cos\,\theta_{x,y}\right] + 4\sqrt{\mathrm{p}_{0,1}}\left[\sqrt{\langle x|x\rangle}\,\cos\,\theta_{x,f} + \sqrt{\langle y|y\rangle}\,\cos\,\theta_{y,f}\right]}_{\gamma} = 0,\end{aligned} \tag{45}$$

where $\cos\theta_{x,z} \equiv |\cos\chi_{x,z}|\cos\varphi_{x,z}$ and analogously for $\cos\theta_{y,z}$, $\cos\theta_{f,z}$, etc. Again, solving the above quadratic equation, we can obtain the positive root of $\sqrt{\langle z|z\rangle}$.

E Parameter estimation

Here, we briefly explain how to estimate the relevant probabilities, $\mathrm{p}_{0,0}, \mathrm{p}_{1,1}$ and $\mathrm{p}_{0,1}$, to compute $S(A|C)$ in Equation (36), to eventually obtain the secret key rate given by Equation (1) from the main text.

Due to the nature of this protocol, in the ideal case, one expects $\mathrm{p}(1) = 1/8$ (see Appendix C for details), which is further reduced in the experimental case of imperfect detectors, etc. Therefore, it is useful if these probabilities could be computed without sacrificing any key-generation rounds. Below, we discuss the case with direct estimation where Alice and Bob use part of the key to obtain these probabilities, as well as the case of indirect estimation where no key-generation rounds are wasted.

E.1 Direct estimation

Here, we sacrifice μ instances of the total N_{raw} key-generation rounds, to directly compute the relevant probabilities. However, since Alice's and Bob's detectors are imperfect, they cannot compute $\mathrm{p}_{0,0} = \mathrm{p}(D_v,R\,;1)$ and $\mathrm{p}_{1,1} = \mathrm{p}(R,D_v\,;1)$ directly, as they cannot differentiate the event D_v,R from the events D_ℓ,R, D'_ℓ,R and $D_\ell D'_\ell,R$, and analogously for R,D_v. However, they can obtain $\tilde{\mathrm{p}}_{0,0} = \mathrm{p}_{0,0} + \mathrm{p}^1_{0,0} + \mathrm{p}^2_{0,0} + \mathrm{p}^3_{0,0} = \mathrm{p}(D_v,R \vee D_\ell,R \vee D'_\ell,R \vee D_\ell D'_\ell,R\,;1)$ directly, and also $\tilde{\mathrm{p}}_{1,1}$. They can then compute $\mathrm{p}^1_{0,0} = \langle k^1_{0,0}|k^1_{0,0}\rangle$, $\mathrm{p}^2_{0,0} = \langle k^2_{0,0}|k^2_{0,0}\rangle$ and $\mathrm{p}^3_{0,0} = \langle k^3_{0,0}|k^3_{0,0}\rangle$, to eventually obtain $\mathrm{p}_{0,0}$. From Equation (31) one has

$$\begin{aligned}
\mathrm{p}^1_{0,0} &= \mathrm{p}(D_\ell,R\,;1) = \frac{\mathrm{p}^A_\ell}{4}\left(\|\,|f^{1,0}_{0,0}\rangle + |f^{1,1'}_{0,1'}\rangle\,\|^2\right),\\
\mathrm{p}^2_{0,0} &= \mathrm{p}(D'_\ell,R\,;1) = \frac{\mathrm{p}^A_\ell}{4}\,\langle f^{1',1}_{0,1}|f^{1',1}_{0,1}\rangle\,,\\
\mathrm{p}^3_{0,0} &= \mathrm{p}(D_\ell D'_\ell,R\,;1) = \frac{\mathrm{p}_\ell^{A^2}}{4}\,\langle f^{2,0}_{0,0}|f^{2,0}_{0,0}\rangle\,.
\end{aligned} \tag{46}$$

Even though Alice and Bob cannot compute the above probabilities, they can estimate them by looking at the events corresponding to the clicks, using the expressions

$$\begin{aligned}
\mathrm{p}(D_c,R\,;1) &= \frac{\mathrm{p}^A_d}{4}\left(\|\,|f^{1,0}_{0,0}\rangle + |f^{1,1'}_{0,1'}\rangle\,\|^2\right),\\
\mathrm{p}(D'_c,R\,;1) &= \frac{\mathrm{p}^A_d}{4}\,\langle f^{1',1}_{0,1}|f^{1',1}_{0,1}\rangle\,,\\
\mathrm{p}(D_cD'_c,R\,;1) &= \frac{\mathrm{p}_d^{A^2}}{4}\,\langle f^{2,0}_{0,0}|f^{2,0}_{0,0}\rangle\,.
\end{aligned} \tag{47}$$

Therefore, we can write $\left(\mathrm{p}^1_{0,0} + \mathrm{p}^2_{0,0}\right)$ and $\mathrm{p}^3_{0,0}$ as

$$\mathrm{p}^1_{0,0} + \mathrm{p}^2_{0,0} = \left(\frac{\mathrm{p}^A_\ell}{\mathrm{p}^A_d}\right)\mathrm{p}(D_c,R \vee D'_c,R\,;1), \qquad \mathrm{p}^3_{0,0} = \left(\frac{\mathrm{p}^A_\ell}{\mathrm{p}^A_d}\right)^2\mathrm{p}(D_cD'_c,R\,;1), \tag{48}$$

where only $\mathrm{p}(D_cD'_c,R\,;1)$ can be obtained using the rounds when Alice gets double clicks in her detector. However, $\mathrm{p}(D_c,R \vee D'_c,R \vee D_\ell D'_c,R \vee D_cD'_\ell,R\,;1)$, corresponding to a single click in Alice's detector, can also be obtained directly. Hence,

$$\mathrm{p}(D_c,R \vee D'_c,R\,;1) = \mathrm{p}(D_c,R \vee D'_c,R \vee D_\ell D'_c,R \vee D_cD'_\ell,R\,;1) - \mathrm{p}(D_\ell D'_c,R\,;1) - \mathrm{p}(D_cD'_\ell,R\,;1). \tag{49}$$

Also, we have

$$\mathrm{p}(D_\ell D'_c,R\,;1) = \frac{\mathrm{p}^A_\ell \mathrm{p}^A_d}{4}\,\langle f^{2,0}_{0,0}|f^{2,0}_{0,0}\rangle = \mathrm{p}(D_cD'_\ell,R\,;1). \tag{50}$$

Therefore, the required probabilities $\mathrm{p}_{0,0}$ and $\mathrm{p}_{1,1}$ are

$$\begin{aligned}
\mathrm{p}_{0,0} &= \tilde{\mathrm{p}}_{0,0} - \left(\frac{\mathrm{p}^A_\ell}{\mathrm{p}^A_d}\right)\mathrm{p}(D_c,R \vee D'_c,R \vee D_\ell D'_c,R \vee D_cD'_\ell,R\,;1) + \left(\frac{\mathrm{p}^A_\ell}{\mathrm{p}^A_d}\right)^2\mathrm{p}(D_cD'_c,R\,;1),\\
\mathrm{p}_{1,1} &= \tilde{\mathrm{p}}_{1,1} - \left(\frac{\mathrm{p}^B_\ell}{\mathrm{p}^B_d}\right)\mathrm{p}(R,D_c \vee R,D'_c \vee R,D_\ell D'_c \vee R,D_cD'_\ell\,;1) + \left(\frac{\mathrm{p}^B_\ell}{\mathrm{p}^B_d}\right)^2\mathrm{p}(D_cD'_c,R\,;1).
\end{aligned} \tag{51}$$

Additionally, to compute $\mathrm{p}_{0,1}$, required to estimate $\mathrm{Re}\,\langle k_{0,0}|k_{1,1}\rangle$ from Equation (42), we use $\mathrm{p}_{0,1} = \tilde{\mathrm{p}}_{0,1} - \mathrm{p}^1_{0,1} - \mathrm{p}^2_{0,1} - \mathrm{p}^3_{0,1} - \mathrm{p}^4_{0,1} - \mathrm{p}^5_{0,1} - \mathrm{p}^6_{0,1}$. Again, using straightforward algebra, we have

$$\begin{aligned}
\mathrm{p}_{0,1} = \tilde{\mathrm{p}}_{0,1} &- \left(\frac{\mathrm{p}^A_\ell}{\mathrm{p}^A_d}\right)\mathrm{p}(D_c,D_v \vee D_c,D'_\ell \vee D'_c,D_\ell \vee D_cD'_\ell,D_v \vee D_\ell D'_c,D_v\,;1)\\
&- \left(\frac{\mathrm{p}^B_\ell}{\mathrm{p}^B_d}\right)\mathrm{p}(D_v,D_c \vee D_\ell,D'_c \vee D'_\ell,D_c \vee D_v,D_cD'_\ell \vee D_v,D_\ell D'_c\,;1)\\
&- 3\left(\frac{\mathrm{p}^A_\ell}{\mathrm{p}^A_d}\right)^2\mathrm{p}(D_cD'_c,D_v\,;1) - 3\left(\frac{\mathrm{p}^B_\ell}{\mathrm{p}^B_d}\right)^2\mathrm{p}(D_v,D_cD'_c\,;1) - 3\left(\frac{\mathrm{p}^A_\ell\,\mathrm{p}^B_\ell}{\mathrm{p}^A_d\,\mathrm{p}^B_d}\right)\mathrm{p}(D_c,D'_c \vee D'_c,D_c\,;1).
\end{aligned} \tag{52}$$

Using the direct estimation method to compute all the relevant probabilities, we obtain the secret key rate r (from Equation (1) from the main text) in Figure 8. We consider the implemented number of rounds, 10^5, as a

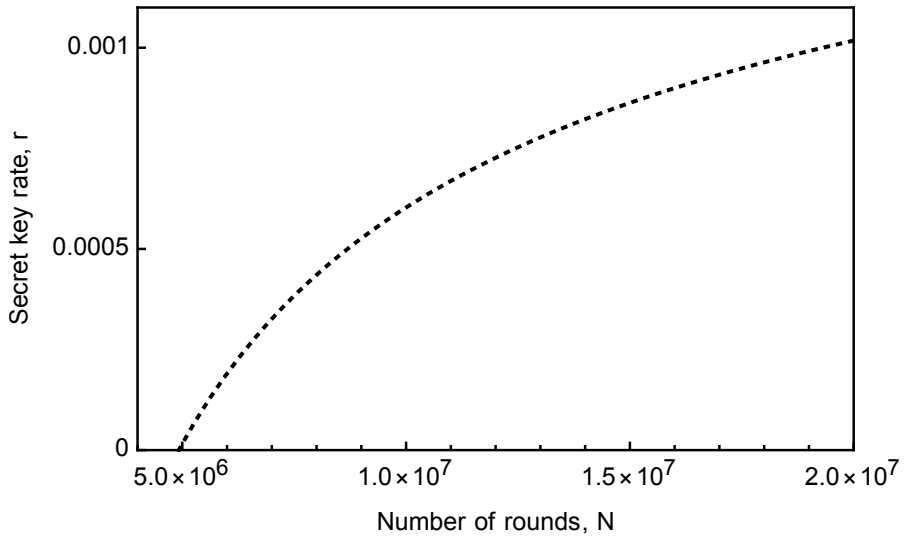


Figure 8: Secret key rate, r, vs number of rounds, N, for the case of imperfect single-photon sources and detectors. The probability necessary for the plot are obtained from the experimental data.

subset of a larger implementation and, therefore, use them to estimate the secret key rate. The probability of server announcing "1" during these rounds is $\mathrm{p}(1) = 0.0162$. Therefore, the amount of keys wasted during the parameter estimations is 1620 bits.

The probabilities of Equations (51) and (52) are the following: $\mathrm{p}_{0,0} = (7.3 \pm 0.3) \times 10^{-3}$, $\mathrm{p}_{1,1} = (5.5 \pm 0.3) \times 10^{-3}$, $\mathrm{p}_{0,1} = (1.1 \pm 0.9) \times 10^{-4}$ and $\mathrm{p}_{1,0} = (5.1 \pm 0.7) \times 10^{-4}$.

We assume $\epsilon = 10^{-5}$, $\epsilon_{EC} = 10^{-10}$ and $\epsilon_{PE} = 10^{-11}$. The value ϵ' is a factor in the min-entropy expression used for the key rate computation and may actually be set by the user arbitrarily to maximize the key rate (see Lemma 1 from [31]). However, for our evaluations we simply set $\epsilon' = 10^{-7}$ (optimizing this could only improve our results). For parameter estimation, we take $\epsilon_{PE} = 10^{-11}$ and assume a confidence interval $\delta = 10^{-4}$, given our experimental errors. The calculated secret key rate corresponds to the minimum lower bound of the entropy $S(A|C)$ (see Equation (36)) over the confidence interval of the experimental probabilities. This minimum occurs for the highest value of the error probability p_{err} and the lowest of p_{key}, and therefore represents the worst possible key rate within our experimental uncertainty.

E.2 Indirect estimation

To avoid wasting the rounds used for key-generation (when "1" was announced without any clicks at Alice's and Bob's detectors), we can use the remaining rounds (when "0","v" or "m" was announced or "1" was announced with click(s) at Alice's and Bob's detectors) for parameter estimation. For these cases, Alice and Bob can communicate over an authenticated channel to convey their respective action choices and resulting states to each other. Therefore, they can communicate for the non-useful rounds where server announces "0","v" or "m", as well as the rounds where any of them detects a photon in case the server announces "1". This method can be applied also in the ideal case described in Section C, but we present it only once for brevity.

We know that $\mathrm{p}_{0,0} = \mathrm{p}(D_v,R\,;1) = \mathrm{p}(D_v,R) - \mathrm{p}(D_v,R\,;0) - \mathrm{p}(D_v,R\,;v) - \mathrm{p}(D_v,R\,;m)$, where $\mathrm{p}(D_v,R) = \mathrm{p}(D,R) - \mathrm{p}(D_\ell,R) - \mathrm{p}(D'_\ell,R) - \mathrm{p}(D_c,R) - \mathrm{p}(D'_c,R) - \mathrm{p}(D_\ell D'_\ell,R) - \mathrm{p}(D_c D'_\ell,R) - \mathrm{p}(D_\ell D'_c,R) - \mathrm{p}(D_c D'_c,R)$. Note that $\mathrm{p}(D,R)$ is the probability of Alice choosing to detect and Bob to reflect. Since Alice and Bob choose their actions at random, ideally $\mathrm{p}(D,D) = \mathrm{p}(D,R) = \mathrm{p}(R,D) = \mathrm{p}(R,R) = 1/4$. However, considering the finite sample size and the inefficiency of switching between the actions, Alice and Bob do not take these probabilities to be $1/4$ but compute them considering only the non-useful rounds. Therefore, we have

$$\begin{aligned}\mathrm{p}_{0,0} = {} & \mathrm{p}(D,R) - \mathrm{p}(D_\ell,R) - \mathrm{p}(D'_\ell,R) - \mathrm{p}(D_c,R) - \mathrm{p}(D'_c,R) - \mathrm{p}(D_\ell D'_\ell,R) - \mathrm{p}(D_c D'_\ell,R) \\ & - \mathrm{p}(D_\ell D'_c,R) - \mathrm{p}(D_c D'_c,R) - \mathrm{p}(D_v,R\,;0) - \mathrm{p}(D_v,R\,;v) - \mathrm{p}(D_v,R\,;m). \end{aligned} \tag{53}$$

Note that Alice and Bob cannot directly compute all the quantities from the above expression, say, $\mathrm{p}(D_v,R\,;0)$, $\mathrm{p}(D_v,R\,;1)$, etc. They can compute $\mathrm{p}(D_\ell,R)$, $\mathrm{p}(D'_\ell,R)$ and $\mathrm{p}(D_c D'_\ell,R)$ analogously as in the previous subsection, given by Equation (48). However, $\mathrm{p}(D_c,R \vee D'_c,R \vee D_\ell D'_c,R \vee D_c D'_\ell,R)$ can be computed directly. We use $\mathrm{p}(D_v,R \vee D_\ell,R \vee D'_\ell,R \vee D_\ell D'_\ell,R\,;0)$, directly observable, to estimate $\mathrm{p}(D_v,R\,;0)$. Therefore,

$$\mathrm{p}(D_v,R\,;0) = \mathrm{p}(D_v,R \vee D_\ell,R \vee D'_\ell,R \vee D_\ell D'_\ell,R\,;0) - \mathrm{p}(D_\ell,R\,;0) - \mathrm{p}(D'_\ell,R\,;0) - \mathrm{p}(D_\ell D'_\ell,R\,;0), \tag{54}$$

$\mathrm{p}(D_\ell,R\,;0)$, $\mathrm{p}(D'_\ell,R\,;0)$ and $\mathrm{p}(D_cD'_\ell,R\,;0)$, etc., can again be computed in the same way as before. Therefore, the final expressions for $\mathrm{p}_{0,0}$ and $\mathrm{p}_{1,1}$, in terms of probabilities computed indirectly, are

$$\begin{aligned}\mathrm{p}_{0,0}=\ &\mathrm{p}(D,R)-\mathrm{p}(D_c,R\vee D'_c,R\vee D_\ell D'_c,R\vee D_cD'_\ell,R)-\mathrm{p}(D_cD'_c,R)-\mathrm{p}(D_v,R\vee D_\ell,R\vee D'_\ell,R\vee D_\ell D'_\ell,R\,;0)\\&-\mathrm{p}(D_v,R\vee D_\ell,R\vee D'_\ell,R\vee D_\ell D'_\ell,R\,;v)-\mathrm{p}(D_v,R\vee D_\ell,R\vee D'_\ell,R\vee D_\ell D'_\ell,R\,;m)\\&+\left(\frac{\mathrm{p}_\ell^A}{\mathrm{p}_d^A}\right)\left[\mathrm{p}(D_c,R\vee D'_c,R\vee D_\ell D'_c,R\vee D_cD'_\ell,R\,;0)+\mathrm{p}(D_c,R\vee D'_c,R\vee D_\ell D'_c,R\vee D_cD'_\ell,R\,;v)\right.\\&\left.+\mathrm{p}(D_c,R\vee D'_c,R\vee D_\ell D'_c,R\vee D_cD'_\ell,R\,;m)-\mathrm{p}(D_c,R\vee D'_c,R\vee D_\ell D'_c,R\vee D_cD'_\ell,R)\right]\\&-\left(\frac{\mathrm{p}_\ell^A}{\mathrm{p}_d^A}\right)^2\left[\mathrm{p}(D_cD'_c,R\,;0)+\mathrm{p}(D_cD'_c,R\,;v)+\mathrm{p}(D_cD'_c,R\,;m)-\mathrm{p}(D_cD'_c,R)\right],\end{aligned}\tag{55}$$

$$\begin{aligned}\mathrm{p}_{1,1}=\ &\mathrm{p}(R,D)-\mathrm{p}(R,D_c\vee R,D'_c\vee R,D_\ell D'_c\vee R,D_cD'_\ell)-\mathrm{p}(R,D_cD'_c)-\mathrm{p}(R,D_v\vee R,D_\ell\vee R,D'_\ell\vee R,D_\ell D'_\ell\,;0)\\&-\mathrm{p}(R,D_v\vee R,D_\ell\vee R,D'_\ell\vee R,D_\ell D'_\ell\,;v)-\mathrm{p}(R,D_v\vee R,D_\ell\vee R,D'_\ell\vee R,D_\ell D'_\ell\,;m)\\&+\left(\frac{\mathrm{p}_\ell^B}{\mathrm{p}_d^B}\right)\left[\mathrm{p}(R,D_c\vee R,D'_c\vee R,D_\ell D'_c\vee R,D_cD'_\ell\,;0)+\mathrm{p}(R,D_c\vee R,D'_c\vee R,D_\ell D'_c\vee R,D_cD'_\ell\,;v)\right.\\&\left.+\mathrm{p}(R,D_c\vee R,D'_c\vee R,D_\ell D'_c\vee R,D_cD'_\ell\,;m)-\mathrm{p}(R,D_c\vee R,D'_c\vee R,D_\ell D'_c\vee R,D_cD'_\ell)\right]\\&-\left(\frac{\mathrm{p}_\ell^B}{\mathrm{p}_d^B}\right)^2\left[\mathrm{p}(R,D_cD'_c\,;0)+\mathrm{p}(R,D_cD'_c\,;v)+\mathrm{p}(R,D_cD'_c\,;m)-\mathrm{p}(R,D_cD'_c)\right].\end{aligned}\tag{56}$$

We can analogously estimate $\mathrm{p}_{0,1}$ by computing $\tilde{\mathrm{p}}_{1,0}$ as

$$\tilde{\mathrm{p}}_{1,0}=\mathrm{p}(R,R\,;1)=\mathrm{p}(RR)-\mathrm{p}(R,R\,;0)-\mathrm{p}(R,R\,;v)-\mathrm{p}(R,R\,;m).\tag{57}$$

Therefore,

$$\begin{aligned}\mathrm{p}_{0,1}=\ &\mathrm{p}(1)-\tilde{\mathrm{p}}_{0,0}-\tilde{\mathrm{p}}_{1,1}-\tilde{\mathrm{p}}_{1,0}-\left(\frac{\mathrm{p}_\ell^A}{\mathrm{p}_d^A}\right)\mathrm{p}(D_c,D_v\vee D_c,D'_\ell\vee D'_c,D_\ell\vee D_cD'_\ell,D_v\vee D_\ell D'_c,D_v\,;1)\\&-\left(\frac{\mathrm{p}_\ell^B}{\mathrm{p}_d^B}\right)\mathrm{p}(D_v,D_c\vee D_\ell,D'_c\vee D'_\ell,D_c\vee D_v,D_cD'_\ell\vee D_v,D_\ell D'_c\,;1)\\&-3\left(\frac{\mathrm{p}_\ell^A}{\mathrm{p}_d^A}\right)^2\mathrm{p}(D_cD'_c,D_v\,;1)-3\left(\frac{\mathrm{p}_\ell^B}{\mathrm{p}_d^B}\right)^2\mathrm{p}(D_v,D_cD'_c\,;1)-3\left(\frac{\mathrm{p}_\ell^A\,\mathrm{p}_\ell^B}{\mathrm{p}_d^A\,\mathrm{p}_d^B}\right)\mathrm{p}(D_c,D'_c\vee D'_c,D_c\,;1).\end{aligned}\tag{58}$$

Note that, to compute $\mathrm{p}_{key}=\tilde{\mathrm{p}}_{00}+\tilde{\mathrm{p}}_{11}$ using the indirect method, we have

$$\begin{aligned}\tilde{\mathrm{p}}_{0,0}=\ &\mathrm{p}(D,R)-\mathrm{p}(D_c,R\vee D'_c,R\vee D_\ell D'_c,R\vee D_cD'_\ell,R)-\mathrm{p}(D_cD'_c,R)-\mathrm{p}(D_v,R\vee D_\ell,R\vee D'_\ell,R\vee D_\ell D'_\ell,R\,;0)\\&-\mathrm{p}(D_v,R\vee D_\ell,R\vee D'_\ell,R\vee D_\ell D'_\ell,R\,;v)-\mathrm{p}(D_v,R\vee D_\ell,R\vee D'_\ell,R\vee D_\ell D'_\ell,R\,;m),\end{aligned}\tag{59}$$

$$\begin{aligned}\tilde{\mathrm{p}}_{1,1}=\ &\mathrm{p}(R,D)-\mathrm{p}(R,D_c\vee R,D'_c\vee R,D_\ell D'_c\vee R,D_cD'_\ell)-\mathrm{p}(R,D_cD'_c)-\mathrm{p}(R,D_v\vee R,D_\ell\vee R,D'_\ell\vee R,D_\ell D'_\ell\,;0)\\&-\mathrm{p}(R,D_v\vee R,D_\ell\vee R,D'_\ell\vee R,D_\ell D'_\ell\,;v)-\mathrm{p}(R,D_v\vee R,D_\ell\vee R,D'_\ell\vee R,D_\ell D'_\ell\,;m).\end{aligned}\tag{60}$$

From our experimental data, we obtain $\mathrm{p}_{0,0}=(8\pm2)\times10^{-3}$, $\mathrm{p}_{1,1}=(6\pm2)\times10^{-3}$, $\mathrm{p}_{0,1}=(3\pm2)\times10^{-3}$ and $\mathrm{p}_{1,0}=(0.5\pm2)\times10^{-3}$. All these values are compatible with those obtained with the direct estimation within experimental uncertainties, which can be reduced by employing a larger sample and improving the single-photon sources and detectors.

F Dependence on detection efficiency

In this section, we discuss the dependence of the secret key rate on the detection efficiencies of Alice's and Bob's detectors, p_d^A and p_d^B, respectively. Note that, since we only consider the first term in Equation (36) to estimate a bound on $S(A|C)$, we only need to compute the probabilities $\mathrm{p}_{0,0}, \mathrm{p}_{1,1}, \mathrm{p}_{0,1}, \mathrm{p}_{1,0}$ and $\mathrm{p}(1)$. However, $\mathrm{p}_{0,0}, \mathrm{p}_{1,1}, \mathrm{p}_{0,1}, \mathrm{p}_{1,0}$ are independent of p_d^A and p_d^B, and it is only $\mathrm{p}(1)$ that has this dependence. Therefore, using the experimental data corresponding to $\mathrm{p}_\ell^A=1-\mathrm{p}_d^A=0.42$ and $\mathrm{p}_\ell^B=1-\mathrm{p}_d^B=0.42$, we can rewrite $\mathrm{p}(1)$ with

the explicit dependence on the general parameters, $\tilde{\mathrm{p}}_\ell^A$ and $\tilde{\mathrm{p}}_\ell^B$, as

$$\begin{aligned}\mathcal{N}(\tilde{\mathrm{p}}_\ell^A,\tilde{\mathrm{p}}_\ell^B) \;&= \langle k_{0,0}|k_{0,0}\rangle + \left(\frac{\tilde{\mathrm{p}}_\ell^A}{\mathrm{p}_\ell^A}\right)\left(\langle k_{0,0}^1|k_{0,0}^1\rangle + \langle k_{0,0}^2|k_{0,0}^2\rangle\right) + \left(\frac{\tilde{\mathrm{p}}_\ell^A}{\mathrm{p}_\ell^A}\right)^2 \langle k_{0,0}^3|k_{0,0}^3\rangle \\ &+ \langle k_{1,1}|k_{1,1}\rangle + \left(\frac{\tilde{\mathrm{p}}_\ell^B}{\mathrm{p}_\ell^B}\right)\left(\langle k_{1,1}^1|k_{1,1}^1\rangle + \langle k_{1,1}^2|k_{1,1}^2\rangle\right) + \left(\frac{\tilde{\mathrm{p}}_\ell^B}{\mathrm{p}_\ell^B}\right)^2 \langle k_{1,1}^3|k_{1,1}^3\rangle \\ &+ \langle k_{0,1}|k_{0,1}\rangle + \langle k_{1,0}|k_{1,0}\rangle + \left(\frac{\tilde{\mathrm{p}}_\ell^A}{\mathrm{p}_\ell^A}\right)\langle k_{0,1}^1|k_{0,1}^1\rangle + \left(\frac{\tilde{\mathrm{p}}_\ell^B}{\mathrm{p}_\ell^B}\right)\langle k_{0,1}^2|k_{0,1}^2\rangle \\ &+ \left(\frac{\tilde{\mathrm{p}}_\ell^A\,\tilde{\mathrm{p}}_\ell^B}{\mathrm{p}_\ell^A\,\mathrm{p}_\ell^B}\right)\left(\langle k_{0,1}^3|k_{0,1}^3\rangle + \langle k_{0,1}^4|k_{0,1}^4\rangle\right) + \left(\frac{\tilde{\mathrm{p}}_\ell^A}{\mathrm{p}_\ell^A}\right)^2 \langle k_{0,1}^5|k_{0,1}^5\rangle + \left(\frac{\tilde{\mathrm{p}}_\ell^B}{\mathrm{p}_\ell^B}\right)^2 \langle k_{0,1}^6|k_{0,1}^6\rangle \\ &= \mathrm{p}(1)(\tilde{\mathrm{p}}_\ell^A,\tilde{\mathrm{p}}_\ell^B).\end{aligned} \tag{61}$$

Moreover, $\mathrm{p}_{err} = \mathrm{p}(1) - \mathrm{p}_{key}$ is also modified accordingly, to be used in computing $Q = \mathrm{p}_{err}/\mathrm{p}(1)$ to obtain the secret key rate presented in Figure 3 from the main paper.

G Dependence on transmission loss

In this section, we provide a brief analysis of the dependence of the key rate on the channel losses, for the case of imperfect photon sources and detectors. The channel loss, after photons passing a distance L through a medium described by the absorption coefficient α (in dB/unit distance), is given by $\ell = \alpha L$. In our protocol, the photons are traveling from the server to the agents, and back, meaning that the total distance L is twice the distance between the server and the agents. This is also the maximal distance between Alice and Bob, achieved when the two are at the opposite sides of the server.

First, note that in general, the all-powerful adversary is bounded only by the laws of physics. In particular, it can vary the number of photons in front of Alice's and Bob's labs at will. But such assumption would seem to turn senseless the whole loss analysis. Moreover, the agents can check the photon number statistics in their labs, thus the adversary must keep them at the levels of the honest case. Finally, note that the overall photon-adversary state in front of the agents has the same shape as in the lossless case. Indeed, expression 11 represents the most general photon-adversary state that contains up to two photons, in which the probabilities of having zero, one, or two photons are incorporated in the norms of vectors $|c_{a,b}\rangle_C \in \mathcal{H}_C$.

In our table-top experimental implementation, due to the low transmission loss in air for the considered distance, we can assume that the loss is for all practical purposes zero. Let us fix the source parameters p_1 and p_2 (the probabilities of single- and double-photon emission per pulse, respectively), the detector efficiency p_d (for simplicity, we assume that the agent's detectors have the same efficiency), and take a certain number of rounds N. For that, we can calculate the key rate $r(\ell = 0; N)$, presented in Figure 4 from the main text. We have that $N = N_0 + N_1 + N_2$, where N_i is the number of rounds with $i = 0, 1, 2$ emitted photons.

Given the transmission probability $T(\ell) = e^{-\frac{\ell}{10}}$, one can calculate

$$\begin{aligned}N_0(\ell) &= N_0 + (1-T)N_1 + (1-T)^2 N_2 \\ N_1(\ell) &= TN_1 + 2T(1-T)N_2 \\ N_2(\ell) &= T^2 N_2,\end{aligned} \tag{62}$$

where $N_i(\ell)$ are the expected numbers of rounds with $i = 0, 1, 2$ photons present. Note that $N_0(\ell) + N_1(\ell) + N_2(\ell) = N_0 + N_1 + N_2 = N$.

Consider the number of rounds $N' < N$, for which $N_1(\ell) = \mathrm{p}_1 N'$ and calculate the secret key $r(0; N')$. Then, we have that $r(\ell; N) \geq r(0; N')$, the secret key for N' rounds in the configuration with $L = 0$ is the lower bound of the secret key for N rounds with the loss ℓ. This bound is based on the following two arguments:

1. The vacuum pulses neither contribute to the key generation, nor to eavesdropping (they leak no information to the adversary). Thus, only the numbers of single-photon emissions and double-photon emissions are relevant, i.e., whenever we have the key rate for the number of rounds that involve certain numbers of the single-photon and double-photon emissions, we can take this result as valid for any case of having the same single- and double-photon rounds (provided there are no higher-photon rounds).

2. Given a certain number of rounds, N, as ℓ grows, both $N_1(\ell)$ and $N_2(\ell)$ decrease. But their ratio does not stay the same, i.e., there exists no N', such that both requirements $N_1(\ell) = \mathrm{p}_1 N'$ and $N_2(\ell) = \mathrm{p}_2 N'$ are

satisfied. In other words, the profile of the source changes with ℓ. But, the ratio $N_1(\ell)/N_2(\ell)$ increases: as ℓ grows, there are proportionally more single-photon rounds than double-photon ones, meaning it is more likely that Alice and Bob receive a single photon than two photons. Since double-photon rounds are the ones that, on one side might induce errors in the key, and on the other help the adversary, we actually have that our $r(\ell = 0; N')$ is in fact the lower bound for $r(\ell > 0; N)$.

Thus, having our results $r(N)$ for $\ell = 0$, our key rate as a function of the loss ℓ is given by

$$\tilde{r}(\ell) \equiv r(N') = r\left(\frac{N_1(\ell)}{\mathrm{p}_1}\right), \tag{63}$$

where by $\tilde{r}$ we denote the functional dependence of the key rate on the losses, which is different from the dependence of r on the number of rounds for $\ell = 0$. Using the second line of (62), $N_1 = \mathrm{p}_1 N$ and $N_2 = \mathrm{p}_2 N$, we finally have

$$\tilde{r}(\ell) = r\Big(\big[10^{-\ell/10}\mathrm{p}_1 + 2 \cdot 10^{-\ell/10}\big(1 - 10^{-\ell/10}\big)\mathrm{p}_2\big] N\Big). \tag{64}$$

References

[1] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. volume 560, pages 7–11, 2014. DOI: https://doi.org/10.1016/j.tcs.2014.05.025.

[2] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991. DOI: https://doi.org/10.1103/PhysRevLett.67.661.

[3] Peter W. Shor and John Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85:441–444, Jul 2000. DOI: https://doi.org/10.1103/PhysRevLett.85.441.

[4] Renato Renner, Nicolas Gisin, and Barbara Kraus. Information-theoretic security proof for quantum-key-distribution protocols. *Phys. Rev. A*, 72:012332, Jul 2005. DOI: https://doi.org/10.1103/PhysRevA.72.012332.

[5] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2053):207–235, 2005. DOI: https://doi.org/10.1098/rspa.2004.1372.

[6] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden. Advances in quantum cryptography. *Adv. Opt. Photon.*, 12(4): 1012–1236, Dec 2020. DOI: https://doi.org/10.1364/AOP.361502.

[7] Akshata Shenoy-Hejamadi, Anirban Pathak, and Srikanth Radhakrishna. Quantum cryptography: Key distribution and beyond. *Quanta*, 6(1):1–47, 2017. ISSN 1314-7374. DOI: https://doi.org/10.12743/quanta.v6i1.57.

[8] Mohsen Razavi, Anthony Leverrier, Xiongfeng Ma, Bing Qi, and Zhiliang Yuan. Quantum key distribution and beyond: introduction. *J. Opt. Soc. Am. B*, 36(3):QKD1–QKD2, Mar 2019. DOI: https://doi.org/10.1364/JOSAB.36.00QKD1.

[9] Feihu Xu, Xiongfeng Ma, Qiang Zhang, Hoi-Kwong Lo, and Jian-Wei Pan. Secure quantum key distribution with realistic devices. *Rev. Mod. Phys.*, 92:025002, May 2020. DOI: https://doi.org/10.1103/RevModPhys.92.025002.

[10] Michel Boyer, Dan Kenigsberg, and Tal Mor. Quantum key distribution with classical bob. *Phys. Rev. Lett.*, 99:140501, Oct 2007. DOI: https://doi.org/10.1103/PhysRevLett.99.140501.

[11] Michel Boyer, Ran Gelles, Dan Kenigsberg, and Tal Mor. Semiquantum key distribution. *Phys. Rev. A*, 79:032341, Mar 2009. DOI: https://doi.org/10.1103/PhysRevA.79.032341.

[12] Walter O. Krawec. Mediated semiquantum key distribution. *Phys. Rev. A*, 91:032323, Mar 2015. DOI: https://doi.org/10.1103/PhysRevA.91.032323.

[13] Zhi-Rou Liu and Tzonelih Hwang. Mediated semi-quantum key distribution without invoking quantum measurement. *Annalen der Physik*, 530(4):1700206, 2018. DOI: https://doi.org/10.1002/andp.201700206.

[14] Xiangfu Zou, Zhenbang Rong, and Nan-Run Zhou. Three attacks on the mediated semi-quantum key distribution without invoking quantum measurement. *Annalen der Physik*, 532(8):2000251, 2020. DOI: https://doi.org/10.1002/andp.202000251.

[15] Po-Hua Lin, Chia-Wei Tsai, and Tzonelih Hwang. Mediated semi-quantum key distribution using single photons. *Annalen der Physik*, 531(8):1800347, 2019. DOI: https://doi.org/10.1002/andp.201800347.

[16] Lingli Chen, Qin Li, Chengdong Liu, Yu Peng, and Fang Yu. Efficient mediated semi-quantum key distribution. *Physica A: Statistical Mechanics and its Applications*, 582:126265, 2021. DOI: https://doi.org/10.1016/j.physa.2021.126265.

[17] Walter O Krawec. Multi-mediated semi-quantum key distribution. In *2019 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6. IEEE, 2019. DOI: https://doi.org/10.1109/GCWkshps45667.2019.9024404.

[18] Julia Guskind and Walter O Krawec. Mediated semi-quantum key distribution with improved efficiency. *Quantum Science and Technology*, 7(3):035019, 2022. DOI: https://doi.org/10.1088/2058-9565/ac7412.

[19] Michel Boyer, Matty Katz, Rotem Liss, and Tal Mor. Experimentally feasible protocol for semiquantum key distribution. *Phys. Rev. A*, 96:062335, Dec 2017. DOI: https://doi.org/10.1103/PhysRevA.96.062335.

[20] Walter O. Krawec. Practical security of semi-quantum key distribution. In Eric Donkor and Michael Hayduk, editors, *Quantum Information Science, Sensing, and Computation X*, volume 10660, pages 33 – 45. International Society for Optics and Photonics, SPIE, 2018. DOI: https://doi.org/10.1117/12.2303759.

[21] Hasan Iqbal and Walter O. Krawec. Semi-quantum cryptography. *arXiv*, 1910.05368, 2019. DOI: https://doi.org/10.48550/arXiv.1910.05368.

[22] Walter O. Krawec. Security proof of a semi-quantum key distribution protocol. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 686–690, 2015. DOI: https://doi.org/10.1109/ISIT.2015.7282542.

[23] Wei Zhang, Daowen Qiu, and Paulo Mateus. Security of a single-state semi-quantum key distribution protocol. *Quantum Information Processing*, 17(6), 2018. ISSN 1570-0755. DOI: https://doi.org/10.1007/s11128-018-1904-z.

[24] R. H. Dicke. Interaction-free quantum measurements: A paradox? *American Journal of Physics*, 49(10): 925–930, 1981. DOI: https://doi.org/10.1119/1.12592.

[25] Avshalom C. Elitzur and Lev Vaidman. Quantum mechanical interaction-free measurements. *Found. Phys.*, 23(7):987–997, Jul 1993. ISSN 1572-9516. DOI: https://doi.org/10.1007/BF00736012.

[26] Paul Kwiat, Harald Weinfurter, Thomas Herzog, Anton Zeilinger, and Mark A. Kasevich. Interaction-free measurement. *Phys. Rev. Lett.*, 74:4763–4766, Jun 1995. DOI: https://doi.org/10.1103/PhysRevLett.74.4763.

[27] Francesco Lenzini, Ben Haylock, Juan C. Loredo, Raphael A. Abrahão, Nor A. Zakaria, Sachin Kasture, Isabelle Sagnes, Aristide Lemaitre, Hoang-Phuong Phan, Dzung Viet Dao, Pascale Senellart, Marcelo P. Almeida, Andrew G. White, and Mirko Lobino. Active demultiplexing of single photons from a solid-state source. *Laser & Photonics Reviews*, 11(3):1600297, 2017. DOI: https://doi.org/10.1002/lpor.201600297.

[28] Leonard Mandel and Emil Wolf. *Optical coherence and quantum optics*. Cambridge university press, 1995.

[29] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov. Invited review article: Single-photon sources and detectors. *Review of Scientific Instruments*, 82(7):071101, 2011. DOI: https://doi.org/10.1063/1.3610677.

[30] Renato Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 6 (01):1–127, 2008. DOI: https://doi.org/10.1142/S0219749908003256.

[31] Valerio Scarani and Renato Renner. Quantum cryptography with finite resources: Unconditional security bound for discrete-variable protocols with one-way postprocessing. *Phys. Rev. Lett.*, 100:200501, May 2008. DOI: https://doi.org/10.1103/PhysRevLett.100.200501.

[32] Walter O. Krawec. Quantum key distribution with mismatched measurements over arbitrary channels. *Quantum Info. Comput.*, 17(3–4):209–241, 2017. ISSN 1533-7146. DOI: https://doi.org/10.26421/QIC17.3-4-2.

[33] Valerio Scarani, Helle Bechmann-Pasquinucci, Nicolas J. Cerf, Miloslav Dušek, Norbert Lütkenhaus, and Momtchil Peev. The security of practical quantum key distribution. *Rev. Mod. Phys.*, 81:1301–1350, Sep 2009. DOI: https://doi.org/10.1103/RevModPhys.81.1301.

[34] Suhri Kim, Sunghyun Jin, Yechan Lee, Byeonggyu Park, Hanbit Kim, and Seokhie Hong. Single trace side channel analysis on quantum key distribution. In *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 736–739, 2018. DOI: https://doi.org/10.1109/ICTC.2018.8539703.

[35] Rupesh Kumar, Francesco Mazzoncini, Hao Qin, and Romain Alleaume. Experimental vulnerability analysis of qkd based on attack ratings. *Scientific Report*, 11(9564), 2021. DOI: https://doi.org/10.1038/s41598-021-87574-4.

[36] Dongjun Park, GyuSang Kim, Donghoe Heo, Suhri Kim, HeeSeok Kim, and Seokhie Hong. Single trace side-channel attack on key reconciliation in quantum key distribution system and its efficient countermeasures. *ICT Express*, 7(1):36–40, 2021. ISSN 2405-9595. DOI: https://doi.org/10.1016/j.icte.2021.01.013.

[37] Shahid Anwar, Zakira Inayat, Mohamad Fadli Zolkipli, Jasni Mohamad Zain, Abdullah Gani, Nor Badrul Anuar, Muhammad Khurram Khan, and Victor Chang. Cross-vm cache-based side channel attacks and proposed prevention mechanisms: A survey. *Journal of Network and Computer Applications*, 93:259–279, 2017. ISSN 1084-8045. DOI: https://doi.org/10.1016/j.jnca.2017.06.001.

[38] Monika Patel, Joseph B. Altepeter, Yu-Ping Huang, Neal N. Oza, and Prem Kumar. Erasing quantum distinguishability via single-mode filtering. *Phys. Rev. A*, 86:033809, Sep 2012. DOI: https://doi.org/10.1103/PhysRevA.86.033809.

[39] Nino Walenta, Tommaso Lunghi, Olivier Guinnard, Raphael Houlmann, Hugo Zbinden, and Nicolas Gisin. Sine gating detector with simple filtering for low-noise infra-red single photon detection at room temperature. *Journal of Applied Physics*, 112(6):063106, 2012. DOI: https://doi.org/10.1063/1.4749802.

[40] W. J. Zhang, X. Y. Yang, H. Li, L. X. You, C. L. Lv, L. Zhang, C. J. Zhang, X. Y. Liu, Z. Wang, and X. M. Xie. Fiber-coupled superconducting nanowire single-photon detectors integrated with a band-pass filter on the fiber end-face. *Superconductor Science and Technology*, 31(3):035012, feb 2018. DOI: https://doi.org/10.1088/1361-6668/aaa6b4.

[41] S. Gao, O. Lazo-Arjona, B. Brecht, K. T. Kaczmarek, S. E. Thomas, J. Nunn, P. M. Ledingham, D. J. Saunders, and I. A. Walmsley. Optimal coherent filtering for single noisy photons. *Phys. Rev. Lett.*, 123: 213604, Nov 2019. DOI: https://doi.org/10.1103/PhysRevLett.123.213604.

[42] Hoi-Kwong Lo, Marcos Curty, and Bing Qi. Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.*, 108:130503, Mar 2012. DOI: https://doi.org/10.1103/PhysRevLett.108.130503.

[43] Kejin Wei, Wei Li, Hao Tan, Yang Li, Hao Min, Wei-Jun Zhang, Hao Li, Lixing You, Zhen Wang, Xiao Jiang, Teng-Yun Chen, Sheng-Kai Liao, Cheng-Zhi Peng, Feihu Xu, and Jian-Wei Pan. High-speed measurement-device-independent quantum key distribution with integrated silicon photonics. *Phys. Rev. X*, 10:031030, Aug 2020. DOI: https://doi.org/10.1103/PhysRevX.10.031030.

[44] Xiaoqing Zhong, Jianyong Hu, Marcos Curty, Li Qian, and Hoi-Kwong Lo. Proof-of-principle experimental demonstration of twin-field type quantum key distribution. *Phys. Rev. Lett.*, 123:100506, Sep 2019. DOI: https://doi.org/10.1103/PhysRevLett.123.100506.

[45] Tae-Gon Noh. Counterfactual quantum cryptography. *Phys. Rev. Lett.*, 103:230501, Dec 2009. DOI: https://doi.org/10.1103/PhysRevLett.103.230501.

[46] Yang Liu, Lei Ju, Xiao-Lei Liang, Shi-Biao Tang, Guo-Liang Shen Tu, Lei Zhou, Cheng-Zhi Peng, Kai Chen, Teng-Yun Chen, Zeng-Bing Chen, and Jian-Wei Pan. Experimental demonstration of counterfactual quantum communication. *Phys. Rev. Lett.*, 109:030501, Jul 2012. DOI: https://doi.org/10.1103/PhysRevLett.109.030501.

[47] G. Brida, A. Cavanna, I.P. Degiovanni, M. Genovese, and P. Traina. Experimental realization of counterfactual quantum cryptography. *Laser Physics Letters*, 9(3):247–252, jan 2012. DOI: https://doi.org/10.1002/lapl.201110120.

[48] Yang Liu, Lei Ju, Xiao-Lei Liang, Shi-Biao Tang, Guo-Liang Shen Tu, Lei Zhou, Cheng-Zhi Peng, Kai Chen, Teng-Yun Chen, Zeng-Bing Chen, and Jian-Wei Pan. Experimental demonstration of counterfactual quantum communication. *Phys. Rev. Lett.*, 109:030501, Jul 2012. DOI: https://doi.org/10.1103/PhysRevLett.109.030501.

[49] Yuan Cao, Yu-Huai Li, Zhu Cao, Juan Yin, Yu-Ao Chen, Hua-Lei Yin, Teng-Yun Chen, Xiongfeng Ma, Cheng-Zhi Peng, and Jian-Wei Pan. Direct counterfactual communication via quantum zeno effect. *Proceedings of the National Academy of Sciences*, 114(19):4920–4924, 2017. DOI: https://doi.org/10.1073/pnas.1614560114.

[50] F. Del Santo and B. Dakić. Two-way communication with a single quantum particle. *Phys. Rev. Lett.*, 120:060503, Feb 2018. DOI: https://doi.org/10.1103/PhysRevLett.120.060503.

[51] Francesco Massa, Amir Moqanaki, Ämin Baumeler, Flavio Del Santo, Joshua A. Kettlewell, Borivoje Dakić, and Philip Walther. Experimental two-way communication with one photon. *Advanced Quantum Technologies*, 2(11):1900050, 2019. DOI: https://doi.org/10.1002/qute.201900050.

[52] Pascale Senellart, Glenn Solomon, and Andrew White. High-performance semiconductor quantum-dot single-photon sources. *Nat. Nanotechnol.*, 12(11):1026, 2017. DOI: https://doi.org/10.1038/nnano.2017.218.

[53] Eric A. Dauler, Matthew E. Grein, Andrew J. Kerman, Francesco Marsili, Shigehito Miki, Sae Woo Nam, Matthew D. Shaw, Hirotaka Terai, Varun B. Verma, and Taro Yamashita. Review of superconducting nanowire single-photon detector system design options and demonstrated performance. *Optical Engineering*, 53(8):1 – 13, 2014. DOI: https://doi.org/10.1117/1.OE.53.8.081907.

[54] T Rudolph and L Grover. Quantum searching a classical database (or how we learned to stop worrying and love the bomb). *arXiv*, 0206066:1–3, 2002. DOI: https://doi.org/10.48550/arXiv.quant-ph/0206066.

Causal orders, quantum circuits and spacetime: distinguishing between definite and superposed causal orders

Nikola Paunković[1] and Marko Vojinović[2]

[1]Instituto de Telecomunicações and Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1049-001, Lisboa, Portugal

[2]Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

arXiv:1905.09682v5 [quant-ph] 23 May 2020

We study the notion of causal orders for the cases of (classical and quantum) circuits and spacetime events. We show that every circuit can be immersed into a classical spacetime, preserving the compatibility between the two causal structures. Using the process matrix formalism, we analyse the realisations of the quantum switch using 4 and 3 spacetime events in classical spacetimes with fixed causal orders, and the realisation of a gravitational switch with only 2 spacetime events that features superpositions of different gravitational field configurations and their respective causal orders. We show that the current quantum switch experimental implementations do not feature superpositions of causal orders between spacetime events, and that these superpositions can only occur in the context of superposed gravitational fields. We also discuss a recently introduced operational notion of an event, which does allow for superpositions of respective causal orders in flat spacetime quantum switch implementations. We construct two observables that can distinguish between the quantum switch realisations in classical spacetimes, and gravitational switch implementations in superposed spacetimes. Finally, we discuss our results in the light of the modern relational approach to physics.

1 Introduction

The notion of causality is one of the most prominent in science, and also in philosophy of Nature. Its treatment separates Aristotelian from the modern physics, and its *formal* meaning within the latter is likely to have played a significant role, over the past centuries since Galileo, in forming our current *everyday* understanding of the notion of causality. While in Newtonian physics the cause-effect relations were encompassed by a rather simple linear and absolute time, Einstein's analysis of causal relations was pivotal in the formulation of the theory of relativity. But it was quantum mechanics (QM) that, through the EPR argument [1], further formalised by Bell [2], showed how quantum nonlocality, rooted in the superposition principle of QM, revolutionised our everyday notion of causality. Finally, strong theoretical evidence that, when combining the two fundamental theories of the modern physics, one is to expect explicit dynamical nonlocal effects in quantum gravity (QG), shows that our basic understanding of causality and causal orders might be crucial in the development of new physics.

Recently, causal orders were, mainly within the quantum information community, discussed in the context of controlled operations. In particular, it was argued that the quantum switch, a specific controlled operation introduced in [3], exhibits superpositions of causal orders, not only in the context of quantised gravity, where genuine superpositions of different states of gravity are present, but also in the experimental realisations performed in classical spacetimes with fixed causal structure [4, 5, 6]. Note that the notion of causal order discussed in these papers is *different* from the causal order of the underlying spacetime structure. We discuss in detail the relation between the two.

In this paper, we analyse the notion of causal orders in the context of classical and quantum circuits, and relate it to the spacetime causal structures. We prove that each circuit can be

Nikola Paunković: npaunkov@math.tecnico.ulisboa.pt
Marko Vojinović: vmarko@ipb.ac.rs

realised in a classical spacetime, preserving the fixed causal relations of the former, with respect to the causal relations between spacetime events of the latter (see the next section for the details of the theorem). Further, we analyse possible realisations of the quantum switch, showing that those performed in everyday labs do not feature superpositions of causal orders between spacetime events (consistent with our theorem), but rather standard non-relativistic quantum mechanical (coherent) superpositions of different evolutions of a system. On the other hand, we argue that genuine superpositions of different causal orders are indeed to be expected within the QG scenario, where superpositions of different states of the gravitational field, with their corresponding causal orders, are manifestly allowed (Hardy was one of the first to discuss the notion of superpositions of causal orders in the context of QG [7]). In addition, we explicitly construct two distinct observables that can distinguish between the realisations of the quantum switch in classical spacetimes, and implementations of the gravitational switch in superposed spacetimes. This way, we show that the two notions of causal orders, namely one discussed in [4, 5, 6] and the other discussed in this paper, can be experimentally distinguished, in contrast to the opposite claim present in the literature [4]. Finally, we discuss our results in the context of the relational approach to physics.

The layout of the paper is as follows. In Section 2, we introduce the notion of causal order for circuits, and prove the Theorem of the circuit immersion in classical spacetimes. Section 3 is devoted to the analysis of the quantum switch implementations in classical spacetimes that do not feature superpositions of spacetime causal orders, as well as implementations in the context of QG. In Section 4, we compare the quantum switch implementations discussed, and introduce observables that can distinguish between those that feature superpositions of spacetime causal orders, and those that do not. Section 5 is devoted to the discussion of the superpositions of causal orders in the context of the relational approach to physics. Finally, in Section 6, we present and discuss the results, provide some final remarks, and list possible future research directions.

2 Causal orders

We begin by discussing circuits and their realisations in (classical) spacetimes with well defined fixed causal orders. Given a directional acyclic graph $G = (I, E)$, where I is the set of graph nodes, and $E = \{(u, v) \,|\, u, v \in I\}$ is the set of its directed edges (arrows pointing from u to v representing the *wires* of the circuit), a *circuit* $\mathcal{C}$ over the set of operations $\mathcal{G}$ is a pair $\mathcal{C} = (G, g)$, where the mapping $g : I \to \mathcal{G}$ assigns operations to each node. Depending on the type of the operations from $\mathcal{G}$, we will call the circuit *classical* (if the operations are, say, classical logic gates), or *quantum* (if the operations are, say, unitaries, measurements, etc.).

The fact that G is directional and acyclic allows one to define a *partial order* $\prec_I$ over the set I as

$$u \prec_I v \quad \stackrel{\text{def}}{\iff} \quad \Big(\exists\, n \in \mathbb{N} \;\wedge\; \{u \equiv u_1, u_2, \ldots, u_n \equiv v\} \subset I\Big) \quad \Big(\forall\, i \in \{1, 2, \ldots, n-1\}\Big)\; (u_i, u_{i+1}) \in E\,, \tag{1}$$

representing the causal relation between the graph nodes. Next, we define the set of *gates of the circuit* $\mathcal{C}$ as $\mathcal{G}_\mathcal{C} = \{g_u \equiv (u, g(u)) \,|\, u \in I\}$. The induced causal order between the circuit gates $\prec_\mathcal{C}$ is by definition given as

$$g_u \prec_\mathcal{C} g_v \quad \stackrel{\text{def}}{\iff} \quad u \prec_I v\,. \tag{2}$$

Moreover, since there exists a canonical bijection between I and $\mathcal{G}_\mathcal{C}$, the order relations $\prec_I$ and $\prec_\mathcal{C}$ are *isomorphic*.

Finally, we can introduce the set $\mathcal{M}$ of all spacetime events, which is assumed to be a traditional $4D$ manifold. On this spacetime manifold we assume to have a gravitational field, described in a standard way, using a metric tensor $g_{\mu\nu}$. The metric is assumed to be of Minkowski signature, such that the metric-induced light cone structure determines a partial order relation between nearby events, denoted $\prec^g_\mathcal{M}$ (or simply $\prec_\mathcal{M}$ when the choice of the metric is implicit). Note that the causal order over the spacetime events is not an intrinsic property of the spacetime manifold itself, but rather determined by the metric, i.e., the configuration of the gravitational field living on the manifold.

One might pose a question if, given a formal circuit $\mathcal{C}$ with gates $\mathcal{G}_\mathcal{C}$, it is possible to realise it in a

lab — if it is possible to "immerse" it into spacetime. More precisely, given an arbitrary spacetime manifold $\mathcal{M}$, our goal is to study if there exists an *order-preserving map* $\mathcal{P} : \mathcal{G}_\mathcal{C} \to \mathcal{M}$, i.e., if the partial order relations satisfy

$$g_u \prec_\mathcal{C} g_v \implies \mathcal{P}(g_u) \prec_\mathcal{M} \mathcal{P}(g_v)\,, \qquad (3)$$

for every $g_u, g_v \in \mathcal{G}_\mathcal{C}$. To that end, we formulate the following theorem (the proof is given in Appendix A).

Theorem. *Any circuit $\mathcal{C}$ can be immersed into a globally hyperbolic spacetime manifold $\mathcal{M}$, such that its relation of partial order $\prec_\mathcal{C}$ is preserved by the relation of spacetime events $\prec_\mathcal{M}$.*

Regarding the physical interpretation of the Theorem, note that it assigns a spacetime point to each gate in a circuit, as opposed to a point in $3D$ space. Since each spatially localised apparatus may perform the same operation more than once, at different moments in time, it may then correspond to several different gates of the circuit, and thus several different nodes of the graph, instead of just one. In other words, a single piece of experimental equipment *does not* always correspond to a single gate of a circuit.

In addition to the above comment, note that in reality each operation actually takes place in some finite volume of both space and time. However, in theoretical arguments it is convenient to approximate this finite spacetime volume with a single point, ignoring the size and time of activity of the device performing the operation. We adopt this approximation throughout this paper.

Circuits are seen as operations acting upon certain inputs to obtain the corresponding outputs. Usually, the initial/final states (which include instructions, measurement results, etc.) are depicted by the wires. But in our approach, the input state is prepared by the "initial gate" $\mathcal{I}$, while the output state is obtained by the "final gate" $\mathcal{F}$. This way, the circuit $\mathcal{C}$ is seen as an operation $\mathcal{O}_\mathcal{C}$ acting from $\mathcal{I}$ to $\mathcal{F}$.

Note that, given a circuit $\mathcal{C}$, the corresponding overall operation $\mathcal{O}_\mathcal{C}$ (as well as the input and the output gates $\mathcal{I}$ and $\mathcal{F}$) is uniquely defined. The opposite is not the case: given the operation $\mathcal{O}$, one can design different circuits $\mathcal{C}, \mathcal{C}', \ldots$ that achieve it. To see this, let us consider the simplest case of the operation which satisfies $\mathcal{O} = \mathcal{O}_2 \circ \mathcal{O}_1$, where $\circ$ represents the composition of operations. This operation can be trivially achieved by the two circuits: (i) $\mathcal{C}$, which consists of three nodes — node i whose gate $\mathcal{I}$ prepares the input state, node o that applies the gate $g_o = \mathcal{O}$, and node f whose gate $\mathcal{F}$ outputs either the quantum state, the classical outcome(s), or the combination of the two; (ii) $\mathcal{C}_{12}$, which consists of four nodes — nodes i and f that perform the same operations as before, and *two* intermediate nodes o_1 and o_2 that perform $g_{o_1} = \mathcal{O}_1$ and $g_{o_2} = \mathcal{O}_2$, respectively. For simplicity, here and elsewhere in the text, by $\mathcal{O}$ we denote both the operation and the gate that implements it. The two situations are depicted in the following diagrams (see Figure 1).

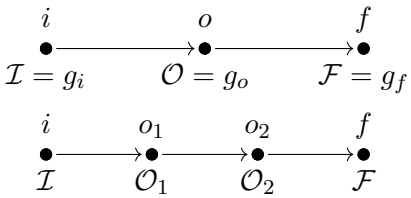


Figure 1: Implementing operation $\mathcal{O}$ with a single gate (upper diagram), and by two consecutive gates $\mathcal{O}_1$ and $\mathcal{O}_2$ (lower diagram).

Finally, in recent literature one can find a notion of an event which is different from the notion of a spacetime point [3, 4, 5, 6, 8, 9, 10, 11]. Namely, one can talk about events as interactions between the quantum system under consideration and the apparatus in the lab. This is motivated by the operational approach to physics, where the interactions between objects are taken as fundamental. Then, one can introduce the relation of partial order, which reflects the causal relationships between such events. Of course, in general, this causal order does not need to coincide with the spacetime causal order. Throughout this paper, if not explicitly stated otherwise, by causal order we mean the order between the spacetime points, which due to our Theorem can also be regarded as the order between the circuit gates. We discuss the difference between the two notions of causal orders in Section 4.

3 Quantum switch

The most prominent feature of quantum systems is that they can be found in *coherent superpositions* of states. This allows for applying the so-

called *control operations*. For simplicity, let us assume that operations $\mathcal{O}$ are unitaries, denoted as U. Given a *control* system C in a superposition $|\varphi\rangle_C = a|0\rangle_C + b|1\rangle_C$ (with $\langle 0|1\rangle_C = 0$), the control operation

$$U_{CT} = |0\rangle_C\langle 0| \otimes U_0 + |1\rangle_C\langle 1| \otimes U_1 \qquad (4)$$

transforms the initial product state $|\Psi_i\rangle_{CT} = |\varphi\rangle_C \otimes |\psi_i\rangle_T$ between the control and the *target* systems into the final entangled state $|\Psi_f\rangle_{CT} = a|0\rangle_C \otimes U_0|\psi_i\rangle_T + b|1\rangle_C \otimes U_1|\psi_i\rangle_T$. A simple realisation of such operation by a circuit consisting of three gates is shown below (see Figure 2).

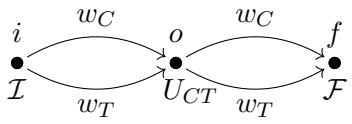


Figure 2: Controlled operation U_{CT}. Applying operation U_b on a system in the wire w_T *controlled* by the state $|b\rangle$ on a system in the wire w_C, with $b = 0, 1$.

Here, the first node and the corresponding gate prepares the initial superposition of the control system, the second implements U_{CT}, and the third is either an identity, a measurement on the two systems, or a combination (say, a measurement of the target qubit, while leaving the control intact). In order to allow for the description of quantum superpositions, we introduce the notion of a vacuum in the analysis of quantum circuits, as is done for example in [12] (for technical details, see Appendix B).

As noted above, given the operation, many different circuits can achieve it. Indeed, in standard optical implementations of the above controlled operation (4), the control qubit is spanned by two spatial modes of a photon, while the target one is its polarisation degree of freedom. The initial superposition state of the control qubit is prepared by a beam splitter, while the two operations U_0 and U_1 are implemented locally in Alice's and Bob's laboratories. Note that, since the control qubit is achieved by the means of two spatial modes of a single photon, while the target qubit is, being the photon's polarisation, "attached to" the control, the target is formally achieved by two degrees of freedom (two wires), one assigned to Alice (T_A), and the other to Bob (T_B). Thus, in such a realization, the control degree of freedom is redundant in the circuit diagram and can be omitted. Nevertheless, since we will later discuss the case of the gravitational quantum switch, in which the gravitational degree of freedom plays the role of the control, here we keep its corresponding wire and gate in the diagram, as presented below (see Figure 3).

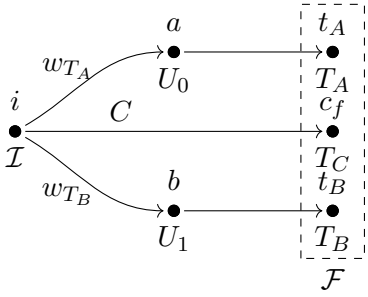


Figure 3: Implementation of the controlled operation using the spatial degree of freedom as a control.

The final gate $\mathcal{F}$ consists of three "elementary gates", represented by the circuit nodes t_A and t_B for the two target wires, and the node c_f for the final control wire.

An important instance of controlled operations is the so-called *quantum switch*, for which the two controlled operations are given by $U_0 = UV$ and $U_1 = VU$, where U and V are two arbitrary unitaries [3]. Having two pairs of equipment, one applying U and the other V, it is straightforward to implement the quantum switch through the circuit similar to the one above, which instead of two gates, one in the node a applying U_0, and another in node b applying U_1, contains four gates placed in the nodes a_U, a_V, b_V and b_U (see Figure 4).

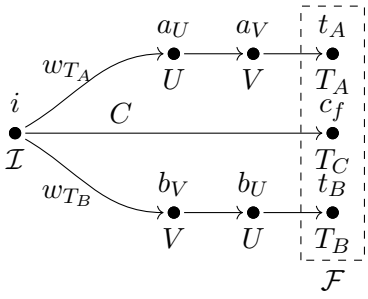


Figure 4: The quantum switch.

The question arises, is it possible to achieve the same using less resources, say, using only two such pieces of equipment, located in two different points (regions) of 3D *space*? Indeed, it is possible

to do so, and recently a number of implementations of the quantum switch were performed in flat Minkowski spacetime [4, 5, 6]. Nevertheless, such implementations still correspond to circuits that implement U_0 and U_1 by four, rather than two gates. The difference is that, when immersing it in a flat spacetime, the two pairs of gates are now distinguished only by the temporal, rather than all four spacetime coordinates. Thus, one cannot talk of superpositions of causal orders between spacetime events in such implementations, as flat (indeed, any globally hyperbolic) spacetime has a manifestly fixed causal order. To implement U_0 and U_1 of the quantum switch by a circuit that consists of two gates only (and thus two corresponding spacetime points), one needs a superposition of gravitational fields with different (incompatible) causal orders. In the following two subsections, we analyse in more detail the "4-event" and the "3-event" implementations of the quantum switch, while the "2-event" case is discussed in the last subsection (the numbers 4, 3 and 2 refer to the numbers of spacetime events corresponding to distinct gates used to achieve U_0 and U_1). A detailed mathematical description using the process matrix formalism [8], is presented in the Appendices C, D and E.

Following the previously mentioned distinction between the spacetime event and the operational notion of the event, the 4-event and 3-event quantum switch implementations will have a description within the operational approach that is different from the spacetime description. In particular, in such approach these two implementations of quantum switch would feature only 2 operationally defined events, and thus the superposition of the corresponding causal orders.

3.1 4-event process

The realisations of the quantum switch are performed in table-top experiments in the gravitational field of the Earth, and can be for all practical purposes considered as being performed in flat Minkowski spacetime. In such experiments, Alice performs the unitary U in her localised laboratory, and Bob performs V in his separate localised laboratory, such that both are stationary with respect to each other and the Earth. The operations are applied on a single particle that arrives from the beam splitter, in a superposition of trajectories towards Alice and Bob, and, upon the exchange between the two agents, is finally recombined on the same beam splitter (for simplicity, we chose one beam splitter, but the whole analysis equally holds for two spatially separated beam splitters), and then measured. Below, we present a spacetime diagram of this experimental realisation of the quantum switch, which also represents a circuit of the implementation scheme (see Figure 5).

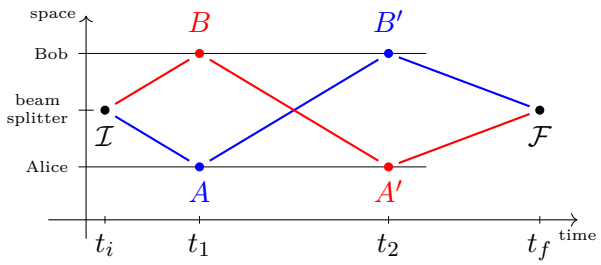


Figure 5: Spacetime diagram, as well as the circuit representation, of the 4-event implementation of the quantum switch.

Black horizontal lines represent world lines for Alice and Bob, as well as the global time coordinate line at the bottom. The black vertical line represents global space coordinate line. Quantum gates are represented by big dots. The composite gate $\mathcal{I}$ consists of the two preparation gates and the initial beam splitter gate, while $\mathcal{F}$ consists of the final beam splitter gate and the target gates that perform the final measurements (for details, see Appendix C). For simplicity, from now on we omit writing the labels of the nodes and keep only the labels of the corresponding circuit gates. The two histories of the particle exchanged between Alice and Bob, representing Alice's and Bob's wires, are full lines coloured in blue and red, respectively.

From the diagram we can see that in the blue history we have the following chain of gates

$$\mathcal{I} \prec_{\mathcal{C}} A \prec_{\mathcal{C}} B' \prec_{\mathcal{C}} \mathcal{F}\,, \tag{5}$$

while for the red history we have

$$\mathcal{I} \prec_{\mathcal{C}} B \prec_{\mathcal{C}} A' \prec_{\mathcal{C}} \mathcal{F}\,. \tag{6}$$

In total, there are four spacetime events involving Alice's and Bob's actions on the particle (gates), namely A, B, A' and B'. Thus, we call the above diagram the "4-event diagram". This setup was already discussed in the literature (see the very end of the Supplementary Notes of [13]).

In order to compare the cases of the quantum and the gravitational switches, it would be interesting to analyse the two examples within recently introduced powerful *process matrix* formalism [8]. To do so, one needs to formulate the formalism involving the vacuum state (see Appendix B for details). The straightforward application of the formalism to the 4-event case is in full accord with the experimental results, as demonstrated in Appendix C.

3.2 3-event process

One can imagine that instead of two, one of the agents implements only one gate. For example, by conveniently choosing the velocity of the particle along its trajectory between Alice and Bob, we can identify Bob's two gates,

$$B \equiv B' \,. \tag{7}$$

We thus arrive to the new spacetime diagram and the associated circuit, called the "3-event diagram" (see Figure 6).

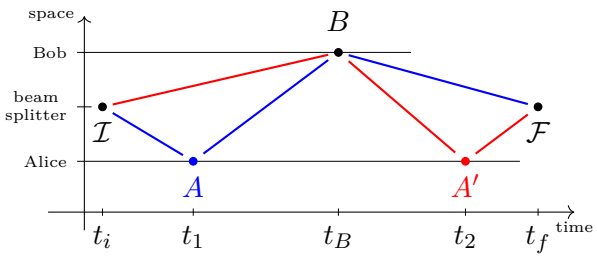


Figure 6: Spacetime diagram, as well as the circuit representation, of the 3-event implementation of the quantum switch.

Now, the obvious question is the following — can we, in addition to (7), impose also that

$$A \equiv A' \,, \tag{8}$$

i.e., also identify Alice's gates into a single spacetime event? In flat Minkowski spacetime, the answer is negative. Namely, by simply looking at the 3-event diagram one can see that the trajectory of the particle between Alice and Bob would require either superluminal speed, or backwards-in-time trajectory in at least one history (note that the diagram assumes that light propagates along the lines that form the 45° angle with the coordinate axes). This is also seen directly from inequalities (5) and (6): identifying both $A \equiv A'$ and $B \equiv B'$ would lead to requiring that *both* $A \prec_{\mathcal{C}} B$ and $B \prec_{\mathcal{C}} A$ are satisfied, i.e., $A \equiv B$. As guaranteed by our Theorem from Section 2, in a curved spacetime it is also impossible to make both identifications (7) and (8), at least if spacetime were globally hyperbolic. Finally, as in the 4-event case, here also the process matrix formalism is consistent with the experimental results, see Appendix D.

3.3 2-event process — gravitational switch

Despite the conclusion of the previous subsection, within the framework of quantum gravity one is allowed to construct superpositions of different gravitational field configurations, leading to superpositions of different causal structures for the spacetime manifold. The assumption of superpositions of different gravitational field configurations is common to all models of QG. Other than that, we will not have any additional assumptions, and thus our approach does not depend on any particular QG model.

In what follows, for the sake of concreteness, we assume the "traditional" approach to the formulation of the QG formalism. Namely, we assume that there exists a smooth $4D$ manifold, called *spacetime*, and denoted as $\mathcal{M}$. Quantum fields, including the gravitational field, live on top of $\mathcal{M}$. The gravitational field is described either via the metric or via some other degrees of freedom (for example, tetrads and spin connection), such that the metric is a function of these. We call this kind of construction "traditional" because it represents a minimal deviation from the mathematical structure of quantum field theory (QFT) in flat Minkowski spacetime, in the sense of preserving the underlying manifold structure. A QG model implementing this approach is, for example, the asymptotic safety framework [14]. Of course, we do not aim to provide a full-fledged model of QG, but rather to only specify the status of the manifold structure within it. As an alternative, in Subsection 5.2, we will discuss the relational framework of QG in which the manifold structure does not exist a priori, but is emergent from relational properties of quantum fields themselves. Finally, note that the discussion of the flat-spacetime cases in the previous sections implicitly assumes the traditional point of view on spacetime manifold. Nevertheless, it has to be compatible with the semiclassical limit of any

viable QG model.

As a consequence of the superposition of causal structures in QG, it is possible to achieve a *gravitational switch*, which implements the same quantum switch as described above, with a circuit consisting (in addition to the initial and final gates $\mathcal{I}$ and $\mathcal{F}$) of only two gates: the Alice's gate A that applies U, and Bob's gate B that applies V. Superposing two gravity-matter states, such that in the first the spacetime geometry (described by the metric tensor g_0) establishes the causal structure

$$\mathcal{I} \prec_{\mathcal{M}}^{g_0} A \prec_{\mathcal{M}}^{g_0} B \prec_{\mathcal{M}}^{g_0} \mathcal{F}\,, \tag{9}$$

while in the second (described by the metric g_1) it is

$$\mathcal{I} \prec_{\mathcal{M}}^{g_1} B \prec_{\mathcal{M}}^{g_1} A \prec_{\mathcal{M}}^{g_1} \mathcal{F}\,, \tag{10}$$

the overall circuit applies operations $U_0 = UV$ and $U_1 = VU$, conditioned on the state of gravity. As a side note, it is clear from (9) and (10) that superpositions of the spacetime causal orders can occur only in the framework of quantum gravity.

Such a switch was previously introduced by Zych *et al.* [15], in the context of two spacetimes which are solutions of the Einstein equations. In their proposal, the beam splitter acted *only* on the gravitational degree of freedom (and the accompanied source, the planet), while leaving the rest of the matter, in particular the particle, Alice and Bob, unaffected. Upon the final beam splitter recombination, the matter is left in an *incoherent* mixture of two states proportional to $\{U\,,V\}|\Psi\rangle$ and $[U\,,V]|\Psi\rangle$. Subsequently, the mass (along with its gravitational degrees of freedom) is being measured in the superposition basis. Upon a post-selection conditioned on the outcome of the measurement, the matter is again in a pure state.

Another way to obtain a genuine superposition of two different causal orders is by using a spatially delocalised beam splitter, that acts on both gravitational and matter fields. This can be depicted by the following 2-event diagram (see Figure 7).

The yellow region in this diagram represents a compact piece of spacetime where the gravitational field is in a superposition of the two distinct states, and plays the role of the control degree of freedom. Along the boundary of that region, both gravitational configurations smoothly join into a single configuration outside. The boundary of the yellow region thus acts as a beam splitter for anything that enters, and again (in the recombining role) for anything that exits. Therefore, all worldlines (namely, of Alice, Bob and the particle) are doubled inside the yellow region. The blue and red colours represent their spacetime trajectories in two different gravitational field backgrounds, respectively.

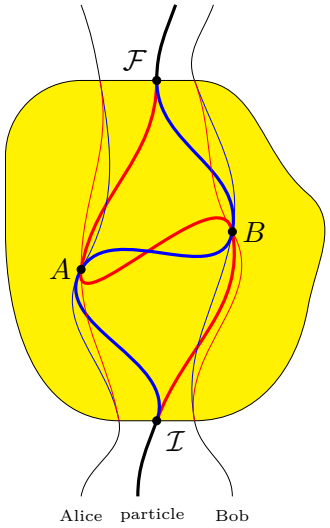


Figure 7: Spacetime diagram of the 2-event implementation of the gravitational switch. Note that formally this is not a circuit diagram, as the control wire, implemented by the state of the gravitational field in the yellow region, is missing.

We model our gravitational switch such that the overall output state is the product between the state of the gravitational field and the state of the particle. The state of the particle is of the form $(\alpha UV + \beta VU)|\Psi\rangle$, obtained without performing final selective measurement. In particular, in order to compare it with the other quantum switch realizations, we choose either $\{U\,,V\}|\Psi\rangle$ or $[U\,,V]|\Psi\rangle$. In order to achieve this, the gravitational switch should act upon *all* degrees of freedom, both gravitational and matter. Note that our gravitational switch does require certain fine tuning, in the sense that the whole, delocalised beam splitter, that acts non-trivially on the whole joint gravity-matter system, is designed for the particular pair of operations applied by Alice and Bob: only for those operations, the beam splitter will output the product state between gravity and matter. Otherwise, the output will be the entangled gravity-matter state, like in the cases of the optical quantum

switch and the gravitational switch introduced by Zych et al. (before the final selective measurement). Still, the process matrix describing the gravitational switch itself is independent of the choice of the gate operations of the agents. See Appendix E for details.

The question whether this kind of diagram is admissible in some theory of quantum gravity is nontrivial, and model dependent, on several grounds. First, it is impossible to construct this diagram by superposing two classical configurations of gravitational field, such that each configuration satisfies Einstein equations. The reason is simple — assuming that the gravitational field is specified outside the yellow region, Einstein equations have a unique solution (up to diffeomorphism symmetry) for the compact yellow region, given such a boundary condition. Therefore, one cannot have two different solutions to superpose inside. The only two options are to either superpose one on-shell and one off-shell configuration of gravity, or two off-shell configurations. This scenario can arguably be considered within the path integral framework for quantum gravity.

Second, the question of the particle trajectory is nontrivial. Namely, given one gravitational configuration in which the particle has the spacetime causal structure (9), corresponding to the blue history, it is not obvious that there can exist another gravitational configuration (with the same boundary conditions at the edge of the yellow region), in which the particle has the spacetime causal structure (10), corresponding to the red history. Even if one admits arbitrary off-shell configurations of gravity, it may turn out that the order of events inside the yellow region must be fixed by the boundary conditions. The only viable way to answer this question is to try and construct an explicit example of two geometries implementing (9) and (10) for the same boundary conditions. Numerical investigations are underway to explore this possibility.

4 Distinguishing 2-, 3-, and 4-event realisations of the quantum switch

In a number of both theoretical proposals [3, 8, 9, 10, 11], as well as experimental realisations [4, 5, 6] of the quantum switch, it is claimed that they feature genuine superpositions of causal orders. The reason for this is the introduction of an alternative, operational notion of the event, which differs from a spacetime point. The motivation for this lies in the claim that the individual spacetime points A and A' (and B and B') do not have an operational meaning. In words of the authors of [4] (see the Discussion section):

> *"The results of the experiment confirm that such [which way] information is not available anywhere and that the interpretation of the experiment in terms of four, causally-ordered events cannot be given any operational meaning. If, on the other hand, one requires events to be defined operationally, in terms of measurable interactions with physical systems [...], then the experiment should be described in terms of only two events — a single use of each of the two gates."*

While it is obvious that the mentioned which-way information is not available in the quantum switch experiment, in what follows we argue that this does not imply that one cannot give an operational meaning to spacetime points, even in the context of the quantum switch in classical geometries.

Below, we first present a critical analysis of the arguments behind introducing the operational notion of event. Then, we show how one can experimentally, at least in principle, distinguish 2-, 3-, and 4-event realisations of the quantum switch.

It is the operational approach to understanding spacetime, applied within the framework of relationalism (see Section 5 for a detailed discussion of the relation between the two frameworks), that is arguably the main argument for introducing the alternative notion of an event. This new notion of an event gives rise to the superposition of respective causal orders in the realisations of the quantum switch even in classical spacetimes. Assuming that the smooth (classical) spacetime is an emergent phenomenon, in the operational approach one considers "closed laboratories" [8] as the primal entities within which one can locally apply standard quantum mechanics, while their connections form the relations from which the spacetime emerges. Indeed, it seems that the process matrix formalism was developed precisely with this idea in mind: to be a mathematical tool in analysing the emergence of the spacetime

through the relations between the closed laboratories. We would like to note that, as shown in Appendices C, D and E, the mentioned formalism is also fully applicable within the standard formulation of quantum mechanics in classical Minkowski spacetime.

Given that in the case of coherent superpositions of the two paths (a particle first goes to Alice, then to Bob, and vice versa) it is not possible to know which of the two has actually been taken, one may conclude that one cannot distinguish between spacetime events A and A', and that the two are operationally given by the single action of a *spatially* localised laboratory. However, this point of view is at odds with our understanding of the ordinary double slit experiment. Namely, by exchanging the roles of time and space, and following the above logic, applied to the case of the standard double slit experiment, one could analogously conclude that, since in order to obtain the interference pattern at the screen one must not (and thus cannot) learn which slit the particle went through, the two slits represent one and the same operational "lab", and one operational point (region) in space.

Let us explain our argument in slightly more detail. Consider first the optical quantum switch. Here, a particle passes through Alice's lab, described by the two spacetime points, (x_A, t) and (x_A, t'). Any attempt to distinguish the times t and t' at which the particle passes through Alice's lab would destroy the superposition. Consider now the standard double slit experiment. Here, a particle passes through the two slits, described by the two spacetime points, (x_L, t) and (x_R, t). Any attempt to distinguish the positions of the slits x_L and x_R through which the particle passes would destroy the superposition. Note that by exchanging the roles of space and time, the descriptions of the above two situations are essentially identical.

According to the operational approach, as a consequence of the above, one should describe Alice's actions in the optical quantum switch in terms of only one operational event. Thus, analogously, one should also describe the particle passing through the slits in terms of only one operational event. However, such interpretation of the double slit experiment is, to the best of our knowledge, absent from the literature.

Note also that the 3-event realisation of the quantum switch offers a natural alternative interpretation of this phenomenon, as a well known *time double slit experiment* [16]. Indeed, the two events (gates) A and A' play the role of the two time-like slits, while the event (gate) B separates the two in the same way the closed shutter separates the two time-like slits in the time double slit experiment. This comes as no surprise: quantum superpositions are in general accompanied by the interference effects, and the quantum switch is, as already emphasised in Section 3, just another instance of a superposition of two different states of the standard quantum mechanics in Minkowski spacetime.

The operational interpretation of identifying the events A and A' in the current experimental realisations of the quantum switch indeed seems to be a tempting proposal. Nevertheless, we would like to point out that in fact it does not resolve any open problem. In addition, being similar to Mach's ideas, it too may be at odds with the theory of general relativity (GR), see Subsection 5.1 for a detailed discussion.

4.1 Distinguishing by decohering the particle

In the above quote from [4], the authors claim that in order to directly distinguish points A and A' (as well as B and B'), one *must* destroy the superposition in the apparatus. Conversely, being unable to distinguish those points in any experiment that maintains superposition and realises the quantum switch, one cannot give them operational meaning. Therefore, those spacetime points are redundant in the theory, and each pair should be replaced by a single operational event. In this subsection, we discuss this type of an argument. In the next, we give an explicit example of an observable that does distinguish such spacetime points without obtaining the which way information.

Let us study one concrete way of distinguishing the mentioned pairs of points, which decoheres the particle. For simplicity, we will analyse the 4- and 2-event cases only. To this end, we will introduce a third agent, Alice's and Bob's *Friend*. At each run of the quantum switch experiment, Alice will, independently and at random, decide whether just to apply her operation onto the particle, or in addition to that, send a photon to Friend. The same holds for Bob. In 25% of the cases, both agents just perform their respective

operations, thus performing the quantum switch. Next, in the 25% of the cases, both agents decide, in addition to applying their respective operations, to send the photons to Friend, who detects them in his spatially localised lab. The remaining 50% of the cases are essentially the same as the previous ones, so for simplicity we omit their analysis.

First, we present the spacetime diagram of the 4-event quantum switch for the case when the agents decide to send the photons to Friend (see Figure 8).

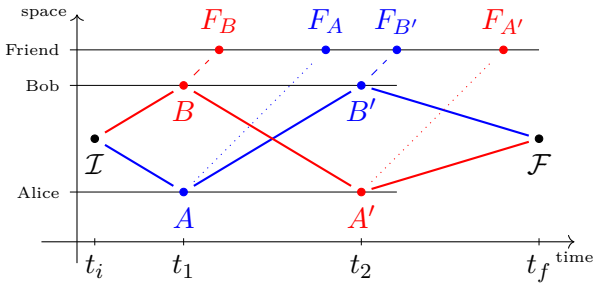


Figure 8: Distinguishing spacetime points by decohering the particle in the 4-event quantum switch. The dotted (dashed) lines represent photons sent by Alice (Bob) to Friend.

The photons coming from Alice are dotted, while the photons coming from Bob are dashed. By knowing the geometry of the whole experiment, Friend would be able to measure, in a generic setup, four *different* times of the photon arrivals: t_A and $t_{A'}$ for spacetime points F_A and $F_{A'}$, and two more for the photons sent by Bob.

On the other hand, in the case of the 2-event gravitational switch realisation, Friend would detect only two times of the photons' arrival. Below, we extend the diagram of the gravitational switch we introduced in Section 3.3, by adding the photons sent to Friend. In order to indicate the fact that the events A and A', etc., are in this setup indeed identified, we write the tilde over the corresponding letters A, B and F (see Figure 9).

Clearly, the two situations are experimentally distinguishable.

Nevertheless, as noted in [4], one might argue that, since the photons sent to Friend in the 4-event case decohere the particle in the switch, this situation does not correspond to the experiment in which the coherence is maintained. Therefore, in the latter, the pair of spacetime events A and

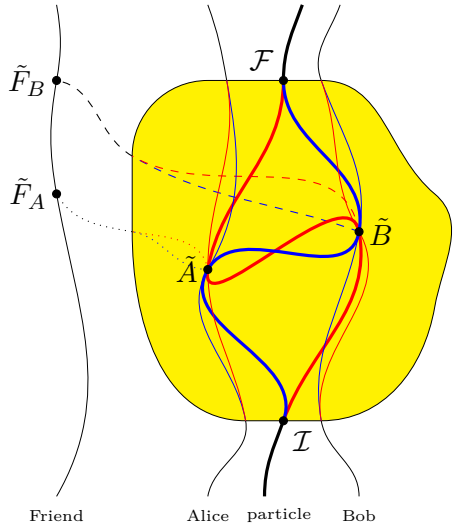


Figure 9: Distinguishing spacetime points by decohering the particle in the 2-event gravitational switch. The dotted (dashed) lines represent photons sent by Alice (Bob) to Friend.

A' still ought to be substituted with a single operational event (and analogously for B and B').

However, even if instead of spacetime points one decides to talk about operational events, such a framework should still be consistent with the experimentally tested theories, GR in particular. According to GR, in flat spacetime (or in any classical configuration of the gravitational field), regardless of whether we decohere the particle or not, both experiments feature four spacetime points, such that A and B (as well as A' and B') can be considered to be simultaneous (see Figure 8). Therefore, the time of execution of both experiments is $\delta t = t_2 - t_1 + C$, where $C \equiv (t_1 - t_i) + (t_f - t_2)$. Note that the time period $t_2 - t_1$ represents the travel time of the particle from one laboratory to the other, and is therefore strictly positive.

From the operational point of view, the decohered version of the experiment also features four operational events, and is thus manifestly consistent with the GR description. Note that a decohered version of the switch still features only two events *per run*: in a classical mixture between "Alice's event before Bob's event" and "Bob's event before Alice's event" each run features just two events, and the duration of the overall experiment in each run is the time between the two events of that run (plus the above constant C).

On the other hand, if the coherence were maintained, the operational point of view features only two operational events, one per laboratory. Then, the total time of execution of the experiment ought to be $\delta\tau = 0 + C$, which is clearly different from the GR prediction. The total time of execution of the quantum switch experiment is a measurable quantity. This means that one can easily determine whether this time is $\delta\tau$ or δt. The former outcome invalidates GR, which would necessitate the formulation of an alternative theory. Note that in this case, a sheer decision to either decohere a particle or not would allow agents to influence the time flow in their labs. Moreover, it raises the question of the time flow in nearby labs *isolated* from the experiment during its execution. The latter outcome poses the problem of the precise formulation of an *operational theory* such that the experiment which features only two operational events lasts precisely the same time as the experiment which features four operational events.

4.2 Distinguishing without decohering the particle

In addition to the above argument, supported by the experimental setup presented in the previous subsection, by erasing the which way information it *is* possible for Friend to distinguish the 4-event and the 2-event realisations even when the "full" quantum switch is executed. For that, one needs to supply Friend with a photon non-demolition measurement. This is in principle possible to construct, although in practice a bit challenging. It thus might be technically easier to use some particles other than photons for sending signals to Friend.

By agreeing in advance of the particular experimental setup, Friend would be able to predict the *distinct* times of arrival of the photons, t_{F_A}, $t_{F_{A'}}$, t_{F_B} and $t_{F_{B'}}$ in the 4-event case, and $t_{\tilde{F}_A}$, $t_{\tilde{F}_B}$ in the 2-event case, thus defining the states of the two photons that arrive to his lab: $|F_A, F_{B'}\rangle$, $|F_{A'}, F_B\rangle$, and $|\tilde{F}_A, \tilde{F}_B\rangle$, respectively. Let us define $\mathcal{H}_{A\prec B'} = \text{span}\{|F_A, F_{B'}\rangle\}$, $\mathcal{H}_{B\prec A'} = \text{span}\{|F_{A'}, F_B\rangle\}$, and $\mathcal{H}_{A\prec B\wedge B\prec A} = \text{span}\{|\tilde{F}_A, \tilde{F}_B\rangle\}$. Then, the relevant Hilbert space of the two photons is

$$\mathcal{H}_{ph} = \mathcal{H}_{A\prec B'} \oplus \mathcal{H}_{B\prec A'} \oplus \mathcal{H}_{A\prec B\wedge B\prec A}\,. \tag{11}$$

Let us define $P_<$, $P_>$ and $P_=$ as orthogonal projectors onto $\mathcal{H}_{A\prec B'}$, $\mathcal{H}_{B\prec A'}$ and $\mathcal{H}_{A\prec B\wedge B\prec A}$, respectively. One can then define a dichotomic photon non-demolition orthogonal observable performed by Friend on the two photons in his laboratory:

$$M = 1\cdot(P_< + P_>) + 0\cdot P_=\,. \tag{12}$$

Provided that the experimental setup is *either* that of the 4-event, or the 2-event type, such measurement would not change the state of the experimental setup (the interferometer, the particle in it, and the photons in the Friend's apparatus), while still leaking the information to Friend (via the measurement outcome) about the type of the quantum switch realisation. Finally, by performing the *quantum erasing* procedure [17, 18], the which way information is lost, and the final state of the particle is restored to a coherent superposition.

Let us examine this more formally. Let the two states of the particle in the quantum switch be $|R\rangle$ and $|B\rangle$, corresponding to the red and the blue trajectory, respectively. After $\mathcal{I}$, the state of the particle in the quantum switch is $\frac{1}{\sqrt{2}}(|R\rangle + |B\rangle)$. As the particle passes through Alice's and Bob's labs, two photons are emitted, which arrive at the Friend's lab. The overall state of the particle and the two photons in the 2-event quantum switch is then

$$\frac{1}{\sqrt{2}}\Big(|R\rangle + |B\rangle\Big)|\tilde{F}_A, \tilde{F}_B\rangle. \tag{13}$$

The particle in the quantum switch is in superposition of the two paths, and it stays so upon measuring M and obtaining the result 0.

On the other hand, the overall state of the particle and the two photons in the 4-event quantum switch is, upon the photons' arrival in the Friend's lab, given by

$$\frac{1}{\sqrt{2}}\Big(|R\rangle|F_{A'}, F_B\rangle + |B\rangle|F_A, F_{B'}\rangle\Big) \tag{14}$$

$$= \frac{1}{2\sqrt{2}}\Big[\Big(|R\rangle + |B\rangle\Big)\Big(|F_{A'}, F_B\rangle + |F_A, F_{B'}\rangle\Big) + \Big(|R\rangle - |B\rangle\Big)\Big(|F_{A'}, F_B\rangle - |F_A, F_{B'}\rangle\Big)\Big].$$

The particle is now decohered by the two photons, and it remains so upon measuring M and obtaining 1 as the result. Therefore, to erase the

which way information, Friend has to perform an additional measurement in the basis

$$|\pm\rangle = \frac{1}{\sqrt{2}}\Big(|F_{A'}, F_B\rangle \pm |F_A, F_{B'}\rangle\Big), \tag{15}$$

thus collapsing the state of the particle in one of the two pure states

$$\frac{1}{\sqrt{2}}\Big(|R\rangle \pm |B\rangle\Big). \tag{16}$$

Knowing the outcome of the measurement of M, Friend can post-select the output of the particle coming out of the quantum switch. Alternatively, in the case of obtaining the $|-\rangle$ result, Friend can change the relative phase between the two of the particle's superposed states.

4.3 Other types of gravitational switches

It is important to note that the framework of QG also allows for the construction of 3- and 4-event switches, in addition to the 2-event one. This is straightforward to see, for example by immersing the above 3- or 4-event spacetime diagram into a superposition of different geometries.

Moreover, all of these gravitational switches may give different outcomes when measuring the observable M, given by (12), followed by the quantum erasing procedure (15). The criteria to necessarily obtain the outcome 0 are: (i) that the photons in red and blue histories meet at the boundary of the yellow region, and (ii) from that point on they recombine into a single photon history. Depending on the details of their construction, all gravitational switches either may or may not satisfy the criteria (i) and (ii). On the other hand, no quantum switch realisations in classical spacetimes with definite causal order could ever yield result 0. Finally, we note that even though some of 2-event gravitational switches may give the outcome 1 when measuring M, it does not necessarily mean that there exist no other observable that could distinguish them from the 4-event quantum switches in a classical geometry. This is a matter for further research.

Detailed graphical visualisations of various gravitational switches are presented in the Appendix F.

5 Relational approach to physics

In the light of the operational framework, which suggests the substitution of the spacetime events A and A' with a single operational event (and analogously for B and B'), it is important to comment on one different but related approach to understanding spacetime, called relationalism. Note that by this promotion of operational events as fundamental entities that ought to replace and play the role of the spacetime events, effectively means the identification of A with A', and B with B'. In this section, we first present a historical review of the relational approach to physics. Then, we discuss the operational framework within the context of the modern approach to relationalism.

5.1 Mach principle and the history of relationalism

The idea of relationalism is an old one, it traces back at least as far as Decartes, and is very important in human thought, in particular in the history of physics. It was brought back to science by Mach in the second half of the XIX century (for an overview and history of the Mach principle and the relational approach to space, from its origins in ancient Greece, see for example [19] and the references therein). Based on the Leibniz ideas of a relational world, Mach formulated his famous Mach principle, an intuitively reasonable approach in analysing physics, and space(time) relations in particular. One of the main characteristics of the Mach principle is that (see [20], page 17):

> *"Space as such plays no role in physics; it is merely an abstraction from the totality of spatial relations between material objects."*

The same formulation can be found in [21], slightly re-phrased as *"Mach7: If you take away all matter, there is no more space."* It is interesting to note that the authors attribute this formulation to A. S. Eddington [22], page 164.

As discussed at the beginning of Section 4, in the operational approach one attributes the ultimate existence to the "closed laboratories" only, while their mutual relations, epitomised by the process matrix, are then giving rise to higher level emergent entities. This clearly shows striking similarities between the Mach's and the operational approaches to space(time).

Mach's ideas were crucial for Einstein in formulating the theory of relativity. And while many of Mach's predictions were indeed realised in the

new theory, some of them were not. Mach's idea that the matter is the basic entity, and that by abstracting the relations between the objects the space emerges, led him to the following statement: if the matter in the universe were finite and had 3D rotational symmetry, it would be impossible to determine its angular momentum (indeed, even talking about it would have no meaning). This is a plausible idea. Nevertheless, it does not hold in general relativity (GR), where one can find two solutions of the Einstein equations for the isolated black hole (the stationary Schwarzschild solution and the rotating Kerr solution [23]). Moreover, while according to the Mach principle the matter completely determines the space, this is not the case in GR: not only that there exists a solution for the gravitational field in the absence of matter (when the stress-energy tensor T is identically zero), but the solution is not unique, as it depends on the boundary conditions as well (i.e., flat Minkowski spacetime is not the only solution — gravitational waves being a possible alternative [23]). This also holds for the general $T \neq 0$ case, as there too boundary conditions play an important role. Thus, matter does not fully determine the inertia, as should according to Mach principle, which states that the inertia of a massive body is given solely in terms of its relations with the other massive bodies.

Motivated by giving the ultimate reality to material objects only (closed laboratories in the case of the operational approach), Mach formulated the above list of claims. Nevertheless, they were later shown not to hold in GR. Provided the similarities between the Mach ideas and the operational approach, the latter might face similar problems as well. We thus believe that introducing the operationalist notion of an event should be accompanied by more elaborate proposals of new physical hypotheses and theories. We hope that our discussion may serve as a small step towards achieving this goal.

5.2 Modern approach to relationalism

In contrast to the historical approach to relationalism and Mach's ideas, that sounded plausible at the time but ultimately failed with the development of GR, the more elaborate modern approach to relationalism is epitomised in the words of Carlo Rovelli (see Section 2.3 of [24]):

> *"The world is made up of fields. Physically, these do not live on spacetime. They live, so to say, on one another. No more fields on spacetime, just fields on fields."*

In particular, the modern relational approach to spacetime defines a particular spacetime point by the physical processes that are "happening at that point". More technically, given an ordered set of classical fields $\phi \equiv (\phi_1, \ldots, \phi_n)$ used to describe physics in a given classical theoretical framework, one traditionally starts from some spacetime point $\tilde{x}$ and evaluates the fields at that point, $\tilde{\phi}_i = \phi_i(\tilde{x})$, obtaining an n-tuple of numbers $(\tilde{\phi}_1, \ldots, \tilde{\phi}_n)$. The idea of relationalism does the opposite — one starts from n-tuples of field values, and then defines a spacetime point using an n-tuple, $\tilde{x} \equiv (\tilde{\phi}_1, \ldots, \tilde{\phi}_n)$, so that the same equation $\tilde{\phi}_i = \phi_i(\tilde{x})$ holds. The question of how to operationally relate values of different fields, and assign and distribute them into n-tuples, is a matter of a separate study [25]. In this work, we assume that this problem is already solved. Moreover, note that fields ϕ need not be observable, due to potential gauge symmetries (for example, the electromagnetic potential A_μ and the metric $g_{\mu\nu}$). To that end, we introduce an ordered set of gauge invariant functions $\mathcal{O}(\phi) \equiv (\mathcal{O}_1(\phi), \ldots, \mathcal{O}_m(\phi))$, where $m \geq n$ (for example, the electromagnetic field strength $F_{\mu\nu}$ and the curvature $R^{\lambda}{}_{\mu\nu\rho}$), and *define a spacetime point* as an m-tuple of their values $\tilde{\mathcal{O}}$.

Unless the physical system features some global symmetry, each m-tuple $\tilde{\mathcal{O}}$ defines a unique point in spacetime. Note that, in the context of GR, the absence of global symmetries is actually the generic case. Thus, the essential feature of this definition is that it does not make sense to say that the same m-tuple of field strengths can occur in two different spacetime points, since "both" spacetime points in question are defined in terms of the one and the same m-tuple, and therefore represent a *single* point.

Moving from classical to the quantum framework, where no system has predetermined physical properties independent of observation, one needs to talk about observables. Given an ordered set of quantum fields $\phi \equiv (\phi_1, \ldots, \phi_n)$, one constructs one *specific* complete set of compatible observables $\mathcal{O} \equiv (\mathcal{O}_1(\phi), \ldots \mathcal{O}_m(\phi))$, where

- *compatible* means that all observables mutu-

ally commute, $[\mathcal{O}_i\,,\,\mathcal{O}_j\,]=0$ for every i and j, while

- *complete* means that the eigenspaces common for all these observables are nondegenerate, i.e., they are one-dimensional subspaces of the total Hilbert space.

Here, by "specific" we mean the set of observables which depend *only* on fields ϕ, but *not* on their conjugated momenta. This fixes the coordinate representation, such that each common eigenvector corresponds to one classical configuration of fields. The outcomes of the measurements of these observables can then be grouped into m-tuples and used to *define individual spacetime points*, as in the classical case above, thus giving rise to an emergent classical spacetime. On the other hand, if the state is not an eigenvector of $\mathcal{O}$, one cannot speak of a single classical configuration of fields, and thus the notion of emergent spacetime and its points ceases to make sense globally, according to the relational approach. At most, one could speak of a superposition of classical configurations and corresponding emergent spacetimes, but without any natural way to relate spacetime points across different branches in the superposition. Nevertheless, this does not mean that establishing such a relation is impossible for certain subregions of spacetime. Indeed, the whole non-yellow "outside" part of the gravitational switch picture from Subsection 3.3 represents a subregion with a locally classical configuration and thus well defined spacetime points.

In order to better appreciate the relational definition of spacetime points given above, it is instructive to look at the realisation of spacetime in the context of a relational quantum gravity model, such as a spinfoam model in the Loop Quantum Gravity (LQG) framework [24, 26]. There, the spacetime is "built" out of the spin foam — a lattice-like structure with vertices, edges and faces, each labeled by the eigenvalues of particular field operators that "live" on these structures, depicted as follows (see Figure 10):

For example, the *area operator*, which is a function of the gravitational field, has eigenvalues determined by a half-integer label $j\in\mathbb{N}/2$, and each face of the spin foam carries one such label, specifying the area of the surface dual to that face. In particular, the spectrum of the area operator

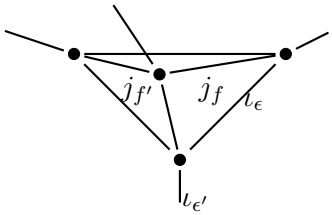


Figure 10: A piece of a spin foam diagram. The field j labels the faces $f, f', \ldots$, while the field ι labels the edges $\epsilon, \epsilon', \ldots$, of the diagram.

is given as

$$A(j)=8\pi\gamma l_p^2\sum_f\sqrt{j_f(j_f+1)}\,, \qquad (17)$$

where l_p is the Planck length, γ is the Barbero-Immirzi parameter, while the sum goes over all faces f of the spin foam that intersect the surface whose area we are interested in, see [24, 26] for details. All other physical observables similarly provide appropriate labels for each vertex, edge and face of the spin foam. Since edges and faces meet at vertices, a given vertex carries labels of all observables of all edges and faces that are connected to that vertex. These observables form the complete set of compatible observables $\mathcal{O}$, and their eigenvalues label each vertex, determining the identity of that vertex. In other words, each labeled vertex of a spin foam defines a "spacetime point", and if two vertices have completely identical properties in the sense of their labels and their connectedness to neighbouring objects, they actually represent the one and the same vertex.

At first sight, it is tempting to apply the ideas of relational spacetime to the case of the quantum switch, as follows. At the spacetime event A, Alice interacts with the particle as it enters and exits her lab, while at the spacetime event A' Alice also interacts (in exactly the same way) with the same particle. The idea of relational spacetime then might suggest that one should *define* the spacetime events A and A' by the physical event of interaction between Alice and the particle. Since this interaction is the same in both cases, one ought to identify the two points, $A\equiv A'$, and claim that both of these correspond to the same spacetime event, defined by the interaction between Alice and the particle. The same argument applies to Bob, and events B and B'.

Unfortunately, this argument is not fully in line

with relationalism. The reason lies in the fact that the interaction between Alice and the particle (and also between Bob and the particle) does not meet the criteria given in the above relational definition of a spacetime point. Namely, neither Alice, nor Bob, performs a measurement of a *complete set of compatible observables* $\mathcal{O}$. The mentioned interaction with the particle is merely a subset of this. In particular, the interaction of Alice with the particle does not uniquely fix the state of, say, the gravitational field, or the electromagnetic field, or the Higgs field, etc. Therefore, it may happen that the measurement outcomes of the whole set of observables $\mathcal{O}$ at spacetime events A and A' are still mutually distinct, thereby defining the events A and A' as two distinguishable spacetime points. In order to be certain that A and A' are really the same spacetime event, Alice would need to measure the complete set of observables $\mathcal{O}$, and convince herself that the results of all those measurements at A and at A' are identical. The mere interaction with the particle is not enough to achieve this, and the experimental setups such as [4, 5, 6] obviously fall short of accounting for the state of all other possible physical fields that Alice and Bob can interact with, in addition to the interaction with the particle.

We see that, when applied to the case of the quantum switch in classical gravitational field, the relational framework is at odds with the operational approach — the former distinguishes A and A' while the latter regards them as identical. This is because the matter fields of the particle are in a superposition of two classical configurations. Similarly, in the case of the 2-event gravitational switch introduced in Subsection 3.3, the overall state of gravity and matter is a superposition of two distinct classical configurations. Therefore, within the relational framework, it is not possible to talk about a single emergent spacetime, nor to compare the points that belong to different branches. This is different from the operational approach, which aims to identify points from different branches. It is also different from the traditional approach, since the latter postulates a unique classical spacetime manifold.

Note that, if understood as an *interpretation*, relational framework ought to have all experimental predictions the same as those from the traditional approach. Thus, the observable constructed in Subsection 4.2 should distinguish the quantum from the gravitational switch, in the same way as in the traditional approach. On the other hand, potential new physics formulated based on the relational framework might, or might not, feature different experimental predictions.

It is important to emphasise that, as discussed in Subsection 4.3, various realisations of the quantum switch are possible by superposing different causal orders in the framework of QG. In particular, regarding the 2-event realisations, one can consider the following diagram (see Figure 11):

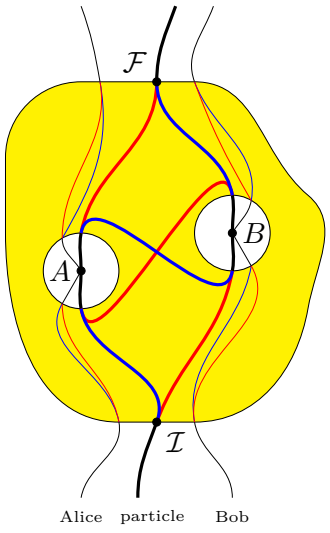


Figure 11: Spacetime diagram of a version of a 2-event gravitational switch, in which Alice and Bob perform their respective operations in the regions of spacetime with a single gravitational configuration.

This diagram features two *classical* spacetime subregions surrounding Alice's and Bob's laboratories. As such, Alice and Bob can measure the complete set of compatible observables within their laboratories, without obtaining which-way information and destroying the superposition. Therefore, even from the relational point of view, this represents an implementation of a 2-event gravitational switch. Note that in this case Alice and Bob do not even need Friend in order to verify the 2-event nature of their gravitational switch.

It is interesting to observe that this realisation of the quantum switch implements the operational idea of a 2-event quantum switch, in terms of closed laboratories. However, to achieve

such an implementation, it is necessary to have a genuine superposition of metric-induced spacetime causal orders in the yellow region of spacetime, which does not feature in experimental realisations [4, 5, 6].

6 Conclusions

In this paper, we analysed the notion of causal orders both in classical and quantum worlds, with the emphasis on the latter. We defined the notion of the causal order for the case of (classical and quantum) circuits, in terms of partial ordering between the nodes of the circuit's underlying graph that defines the cause-effect structure. We discussed the possibility of implementing an abstract circuit in the real world, showing that it is always possible to do so for the case of a globally hyperbolic (classical) spacetime, in which the circuit's causal order is preserved by the metric-induced relation between the spacetime events.

The superposition principle of quantum mechanics offers the possibility of controlled operations, in particular the quantum switch, whose experimental realisations have been claimed to present genuine superpositions of causal orders. Within the process matrix formalism, we have analysed the 4- and 3-event realisations of the quantum switch in classical spacetimes with fixed spacetime causal orders, and the 2-event realisation of a gravitational switch that features superpositions of different gravitational field configurations and their respective spacetime causal orders. To that end, we have extended the process matrix formalism, by introducing the notion of a vacuum state. Our analysis shows that the process matrix formalism can explain the quantum switch realisations within the standard physics, and is thus consistent with it.

Thus, as a consequence of our Theorem, and the analysis of the quantum switch implementations, we argued that, in contrast to the gravitational switch, the current experimental implementations do not feature superpositions of spacetime causal orders, and that they are variants of the time double slit experiment. Moreover, by explicitly constructing two different observables, presented in Sections 4.1 and 4.2, respectively, we showed that it is possible to experimentally distinguish between different realisations of the quantum switch.

Finally, in Section 5, we analysed the relation among the traditional QFT approach to QG (used throughout this paper), the operational point of view, and the relational framework of QG. On the example of the quantum switch, we showed that the operational viewpoint, while consistent with the approach advocated by Mach, is nevertheless at odds with the modern relational framework. On the other hand, the traditional QFT approach and the relational framework may or may not be compatible, depending on the concrete realisation of the quantum switch. In particular, for the specific realisation of the gravitational switch given in Subsection 5.2, the two frameworks are compatible in the prediction that Alice and Bob can locally (without the help of Friend) verify that the switch is implemented on 2 events.

In a recent work [27], the authors report on a violation of the causal inequality [8] in flat Minkowski spacetime with a definite causal order. To achieve it, they consider laboratories that are localised in space only, while delocalised in time. Therefore, their alternative notion of a "closed laboratory", and that considered in [8], do not coincide, this way manifestly violating the conditions necessary for the causal inequality to hold. For the same reason, the scenario considered in [27] falls out of the scope of the current work as well. Additionally, in another recent work [28], the author discusses the quantum switch in terms of the time-delocalised quantum subsystems and operations, and generalises it to more complex quantum circuits and processes. The results of these two papers deserve further analysis and remain to be a subject of future research.

Exploring possible generalisations of our Theorem, as suggested at the end of Appendix A, presents a straightforward future line of research. Also, one could further analyse the process matrix formalism, in particular by exploring the situations in which the operational approach interpretation fails to describe the known processes. Or, to search for the opposite — the instances of physical processes that cannot be explained by the process matrix formalism, when applied within the standard physics. In order to show that the process matrix formalism is perfectly suitable for describing the quantum switch implementations within the standard physics, we

formulated its version that features the vacuum state. One can thus further study possible generalisations of this formalism and its applications to the cases that go beyond simple non-relativistic mechanics. Finally, motivated by our analysis and discussion from Subsections 4.1 and 5.1, one can try to formulate alternative theories that would be consistent with the experimentally tested known physics (GR in particular), while at the same time substituting the spacetime events A and A', from the quantum switch realisations in classical spacetimes, with a single operational event (and analogously for B and B').

Acknowledgments

The authors would like to thank Časlav Brukner, Esteban Castro-Ruiz and Flaminia Giacomini for helpful comments and discussions.

NP acknowledges the support of SQIG — Security and Quantum Information Group, the Instituto de Telecomunicações (IT) Research Unit, ref. UIDB/50008/2020, funded by Fundação para a Ciência e Tecnologia (FCT), and the FCT projects QuantumMining POCI-01-0145-FEDER-031826 and Predict PTDC/CCI-CIF/29877/2017, supported by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalisation Operational Programme (COMPETE 2020), and by the Regional Operational Program of Lisbon, as well as the FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT0006. The bilateral scientific cooperation between Portugal and Serbia through the project "Noise and measurement errors in multi-party quantum security protocols", no. 451-03-01765/2014-09/04 supported by the FCT, Portugal, and the Ministry of Education, Science and Technological Development of the Republic of Serbia is also acknowledged.

MV was supported by the project ON 171031 of the Ministry of Education, Science and Technological Development of the Republic of Serbia, and the bilateral scientific cooperation between Austria and Serbia through the project "Causality in Quantum Mechanics and Quantum Gravity — 2018-2019", no. 451-03-02141/2017-09/02 supported by the Austrian Academy of Sciences (ÖAW), Austria, and the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] A. Einstein, B. Podolsky and N. Rosen, *Physical Review* **47**, 777 (1935).

[2] J. S. Bell, *Physics Physique Fizika* **1**, 195 (1964).

[3] G. Chiribella, G. M. D'Ariano, P. Perinotti and B. Valiron, *Physical Review A* **88**, 022318 (2013).

[4] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner and P. Walther, *Nature Communications* **6**, 7913 (2015).

[5] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner and P. Walther, *Science Advances* **3**, e1602589 (2017).

[6] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner and P. Walther, `arXiv:1712.06884` .

[7] L. Hardy, *Journal of Physics A: Mathematical and Theoretical* **40**, 3081 (2007).

[8] O. Oreshkov, F. Costa and Č. Brukner, *Nature Communications* **3**, 1092 (2012).

[9] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi and Č. Brukner, *New Journal of Physics* **17**, 102001 (2015).

[10] J. Bavaresco, M. Araújo, Č. Brukner and M. T. Quintino, *Quantum* **3** 176 (2019).

[11] P. Allard Guérin, G. Rubino and Č. Brukner, *Phys. Rev. A* **99** 062317 (2019).

[12] C. Portmann, C. Matt, U. Maurer, R. Renner and B. Tackmann, *IEEE Transactions on Information Theory* **63**, 3277 (2017).

[13] J.-P. W. MacLean, K. Ried, R. W. Spekkens and K. J. Resch, *Nature Communications* **8**, 15149 (2017).

[14] M. Niedermaier and M. Reuter, *Living Rev. Rel.* **9**, 5 (2006).

[15] M. Zych, F. Costa, I. Pikovski and Č. Brukner, *Nature Communications* **10**, 3772 (2019).

[16] M. Bauer, *American Journal of Physics* **82**, 1087 (2014).

[17] M. O. Scully and K. Drühl, *Physical Review A* **25**, 2208 (1982).

[18] M. O. Scully, B.-G. Englert and H. Walther, *Nature* **351**, 111 (1991).

[19] H. Lichtenegger and B. Mashhoon, in *The Measurement of Gravitomagnetism: A Challenging Enterprise*, edited by L. Iorio, 13, Nova Science Pub Inc, New York (2007).

[20] M. Blagojević, *Gravitation and Gauge Symmetries*, Institute of Physics Publishing, Bristol (2002).

[21] H. Bondi and J. Samuel, *Physics Letters A* **228**, 121 (1997).

[22] A. S. Eddington, *Space time and gravitation*, Cambridge University Press, Cambridge (1921).

[23] C. Misner, K. Thorne and J. Wheeler, *Gravitation*, W. H. Freeman, San Francisco (1973).

[24] C. Rovelli, *Quantum Gravity*, Cambridge University Press, Cambridge (2004).

[25] J. Janjić, N. Paunković and M. Vojinović, *in preparation*.

[26] C. Rovelli and F. Vidotto, *Covariant Loop Quantum Gravity: An Elementary Introduction to Quantum Gravity and Spinfoam Theory*, Cambridge University Press, Cambridge (2014).

[27] C. T. Marco Ho, F. Costa, C. Giarmatzi and T. C. Ralph, `arXiv:1804.05498` .

[28] O. Oreshkov, *Quantum* **3**, 206 (2019).

[29] R. A. Bertlmann and P. Krammer, *Journal of Physics A: Mathematical and Theoretical* **41**, 235303 (2008).

A Proof of the Theorem

Here we give an explicit constructive proof of the Theorem from the main text.

Given the graph G, we begin the proof by partitioning its set of nodes I into disjoint subsets, in the following way. Since the graph is finite, we introduce the subset $M_1 \subset I$ which consists of all minimal nodes of the graph G:

$$M_1 = \{u \in I \,|\, (\neg\exists v \in I)\, v \prec_I u\}\,. \tag{18}$$

Since all nodes in M_1 are minimal, there is no order relation $\prec_I$ between any two of them. Therefore, we can intuitively understand them as "simultaneous". As a next step, we remove these nodes and the corresponding edges from G, reducing it to a subgraph $G_2 = (I_2, E_2)$, where

$$I_2 = I \backslash N_1\,, \qquad E_2 = \{(u,v) \,|\, u,v \in I_2\,,\, (u,v) \in E\}\,. \tag{19}$$

Then we repeat the construction for the graph G_2, obtaining the new minimal set M_2, and the next subgraph G_3, in an analogous way. Since the graph G is finite, after a certain finite number of steps we will exhaust all nodes in I, ending up with a partition of "simultaneous" subsets $M_1, \ldots, M_m$ ($m \in \mathbb{N}$), such that

$$(\forall i \neq j) \quad M_i \cap M_j = \emptyset\,, \qquad \bigcup_{i=1}^{m} M_i = I\,. \tag{20}$$

Once we have partitioned the set of nodes I into subsets, we turn to the construction of the immersing map $P : I \to \mathcal{M}$, in the following way. Since spacetime is globally hyperbolic, we can write $\mathcal{M} = \Sigma \times \mathbb{R}$, where Σ is a spatial 3-dimensional hypersurface, and $\mathbb{R}$ is timelike. Without loss of generality, one can then introduce a foliation of spacetime into a family of such hypersurfaces, denoted Σ_t and labeled by a parameter $t \in \mathbb{R}$. Start from some initial parameter t_1, and choose a compact subset $S_{t_1} \subset \Sigma_{t_1}$. Denoting the number of elements in the partition M_i as $\|M_i\|$, we pick in an arbitrary way the set of $\|M_1\|$ points $\vec{x}_k \in S_{t_1}$ (here, $k = 1, \ldots, \|M_1\|$), and define the map P to assign a node from M_1 to each point $\vec{x}_k$ in a one-to-one fashion:

$$P(u_k) = (t_1, \vec{x}_k) \in \mathcal{M}\,, \qquad k = 1, \ldots, \|M_1\|\,. \tag{21}$$

Once this assignment has been defined, construct a future-pointing light cone from each spacetime point $(t_1, \vec{x}_k)$. Then we find a new hypersurface, Σ_{t_2}, which contains a common intersection with all constructed light cones, and denote this intersection $S_{t_2} \subset \Sigma_{t_2}$. In this way, by construction, all points $(t_1, \vec{x}_k)$ are in the past of all points in S_{t_2},

$$(t_1, \vec{x}_k) \prec_{\mathcal{M}} S_{t_2}\,, \qquad k = 1, \ldots, \|M_1\|\,. \tag{22}$$

Now extend the definition of P such that it assigns the nodes from the next partition, M_2, to a randomly chosen set of points in S_{t_2} in a similar way as before, then construct a set of light cones from them, and repeat the construction for all partitions M_i. Constructed in this way, the map P ensures that for every pair of nodes $u, v \in I$, we have

$$u \prec_I v \quad \Longrightarrow \quad P(u) \prec_{\mathcal{M}} P(v)\,, \qquad \forall u, v \in I\,. \tag{23}$$

Once we have constructed the map $P : I \to \mathcal{M}$ satisfying (23), using the definition (2), it induces the map $\mathcal{P} : \mathcal{G}_{\mathcal{C}} \to \mathcal{M}$, which satisfies the required statement (3).

This completes the proof. □

Note that, while the causal order $\prec_{\mathcal{M}}$ indeed preserves the causal order $\prec_{\mathcal{C}}$, it is "stronger" in the sense that it may introduce additional relations between the images of nodes, which do not hold in the graph itself. Indeed, the construction of the map P in the above proof is such that *each* image of a node from some given partition M_i is in the causal past of *all* images from the previous partition M_{i-1}, which is not necessarily the case for the nodes themselves. One might study if the causal orders over

the set of nodes and over the set of its images can be equivalent, i.e., if the opposite implication from equation (23) also holds (in this case the immersion P is called an *embedding* of G into $\mathcal{M}$). Whether such an embedding exists for all hyperbolic spacetimes, or at least for some, is an open question.

Next, one could also discuss the generalisation of the above theorem to the case of countably infinite graphs G. However, for our purposes, the existence of a partially ordered map $\mathcal{P}$ over the set of finite graphs will suffice.

Regarding the proof itself, one can formulate an alternative (and simpler) approach to the proof of the theorem. Namely, one can first prove that every circuit can be immersed into the flat Minkowski spacetime. Then, knowing that the sufficiently small neighbourhood of every spacetime point in an arbitrary manifold $\mathcal{M}$ can be well approximated with its tangent space, one can always immerse the whole circuit into this small neighbourhood. However, this implies that the geometric size of the circuit can be considered negligible compared to the curvature scale of the manifold, which may render such implementation practically unfeasible. Moreover, this alternative approach does not cover the cases where one actually wants the scale of the circuit to be comparable to the curvature scale. Specifically, if one wishes to employ the circuit to study gravitational phenomena, its gates must be distributed across spacetime precisely in a way that is sensitive to curvature. Therefore, the construction of the map P used in the proof of the theorem is more general than the construction in this alternative approach.

Finally, given the construction in the proof, the gates of the set of minimal nodes M_1 define the initial gate $\mathcal{I}$, the set of maximal nodes M_m define the final gate $\mathcal{F}$, while the the gates of the remaining intermediary sets of nodes $M_2, \dots M_{m-1}$ define the operation $\mathcal{O}_\mathcal{C}$. This is illustrated in the diagram below (see Figure 12).

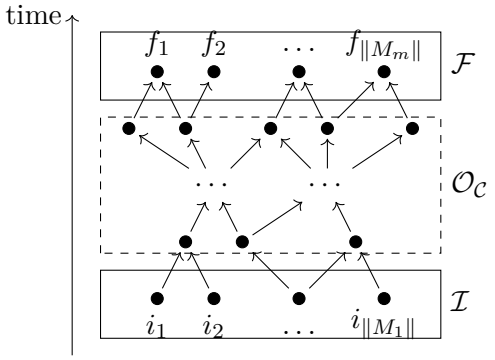


Figure 12: The spacetime diagram of the circuit $\mathcal{C}$, with the initial gate $\mathcal{I}$, the operation gate $\mathcal{O}_\mathcal{C}$, and the final gate $\mathcal{F}$.

B Qutrit states, operators and bases

The notion of a qubit can be generalised from a 2-dimensional Hilbert space to a d-dimensional Hilbert space. The generalised object is called "qudit" in d dimensions [29]. Since we are interested in describing ordinary 2-dimensional qubits with an additional vacuum state, it is natural to consider qudits in $d = 3$, called "qutrits". We introduce the following notation for the basis states of a qutrit in $\mathcal{H}_3 = \mathbb{C}^3$:

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad |1\rangle \equiv \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad |v\rangle \equiv \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{24}$$

The states $|0\rangle$ and $|1\rangle$ will be understood as the usual computational basis for a 2-dimensional qubit, while the state $|v\rangle$ will represent the vacuum, i.e., the "absence of a qubit". In cases when we take sums over the basis vectors, we will assume that the vacuum state carries the index 2, i.e., $|v\rangle \equiv |2\rangle$,

so that we can write

$$\sum_{i=0}^{2} |i\rangle = |0\rangle + |1\rangle + |v\rangle\,, \qquad \text{and} \qquad \sum_{i=0}^{1} |i\rangle = |0\rangle + |1\rangle\,. \tag{25}$$

Using this notation, we write the unnormalised maximally correlated states for the qutrit and the qubit as

$$|\mathbb{1}\rangle\!\rangle = \sum_{i=0}^{2} |i\rangle|i\rangle = |0\rangle|0\rangle + |1\rangle|1\rangle + |v\rangle|v\rangle \in \mathcal{H}_3 \otimes \mathcal{H}_3\,, \qquad |1\rangle\!\rangle = \sum_{i=0}^{1} |i\rangle|i\rangle = |0\rangle|0\rangle + |1\rangle|1\rangle \in \mathcal{H}_2 \otimes \mathcal{H}_2\,, \tag{26}$$

so that

$$|\mathbb{1}\rangle\!\rangle = |1\rangle\!\rangle + |v\rangle|v\rangle\,. \tag{27}$$

One can also introduce the standard Hilbert-Schmidt basis in the space $\mathcal{L}(\mathcal{H}_3)$ of linear operators on $\mathcal{H}_3$. This basis consists of 9 matrices 3×3, labeled as $\lambda_0, \ldots, \lambda_8$, as follows:

- the three symmetric matrices

$$\lambda_1 = \sqrt{\frac{3}{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \lambda_2 = \sqrt{\frac{3}{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \lambda_3 = \sqrt{\frac{3}{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \tag{28}$$

- the three antisymmetric matrices

$$\lambda_4 = \sqrt{\frac{3}{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \lambda_5 = \sqrt{\frac{3}{2}} \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \qquad \lambda_6 = \sqrt{\frac{3}{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \tag{29}$$

- and the three diagonal matrices

$$\lambda_7 = \sqrt{\frac{3}{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \qquad \lambda_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{30}$$

The matrix λ_0 is the unit matrix, while $\lambda_1, \ldots, \lambda_8$ are self-adjoint, traceless, and orthogonal with respect to the standard scalar product:

$$\lambda_i^\dagger = \lambda_i\,, \qquad \operatorname{Tr} \lambda_i = 0\,, \qquad \operatorname{Tr} \lambda_i^\dagger \lambda_j = 3\delta_{ij}\,, \qquad i = 1, \ldots, 8\,. \tag{31}$$

They represent the generators of the $SU(3)$ group, and are known as the Gell-Mann matrices (up to a normalisation factor $\sqrt{3/2}$).

If we denote $\mathcal{H}_v$ as the 1-dimensional vacuum-spanned subspace of $\mathcal{H}_3$, one can see that $\mathcal{L}(\mathcal{H}_2) \oplus \mathcal{L}(\mathcal{H}_v) \subset \mathcal{L}(\mathcal{H}_3)$. In particular, if we denote the standard Pauli matrices as $\sigma_x, \sigma_y, \sigma_z$ and the unit 2×2 matrix as I_2, they form the basis in $\mathcal{L}(\mathcal{H}_2)$, and the qubit basis can thus be constructed as

$$\sqrt{\frac{2}{3}}\lambda_1 = \left[\begin{array}{cc|c} & \sigma_x & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline 0 & 0 & 0 \end{array}\right], \qquad \sqrt{\frac{2}{3}}\lambda_4 = \left[\begin{array}{cc|c} & \sigma_y & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline 0 & 0 & 0 \end{array}\right], \qquad \sqrt{\frac{2}{3}}\lambda_7 = \left[\begin{array}{cc|c} & \sigma_z & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline 0 & 0 & 0 \end{array}\right], \tag{32}$$

along with

$$\frac{2}{3}\lambda_0 + \frac{\sqrt{2}}{3}\lambda_8 = \left[\begin{array}{cc|c} & I_2 & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline 0 & 0 & 0 \end{array}\right]. \tag{33}$$

Also, the vacuum space $\mathcal{L}(\mathcal{H}_v)$ is one-dimensional, and the basis is

$$\frac{1}{3}\lambda_0 - \frac{\sqrt{2}}{3}\lambda_8 = \left[\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 1 \end{array}\right]. \tag{34}$$

C Process matrix evaluation

Let us give an explicit step by step evaluation of the probability distribution for the 4-event process discussed in the text, using the process matrix formalism. The complete spacetime diagram of the process is given as (see Figure 13):

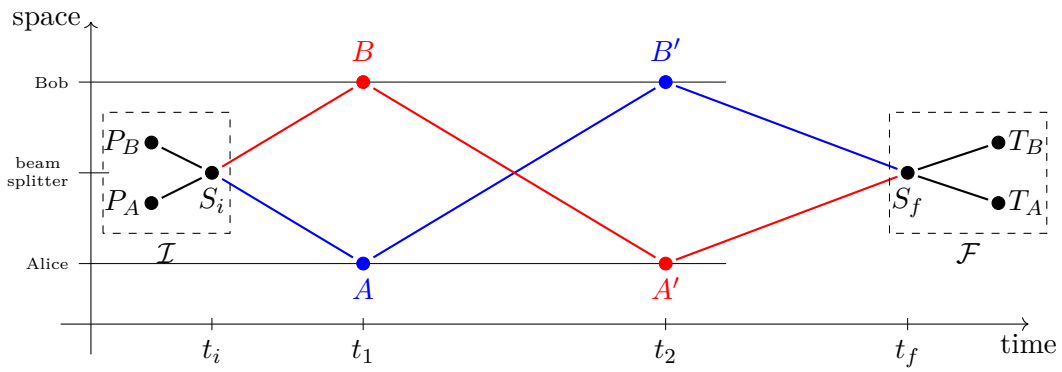


Figure 13: Spacetime diagram of the 4-event implementation of the quantum switch. The internal structures of the composite gates $\mathcal{I}$ and $\mathcal{F}$ are explicitly depicted.

The composite event $\mathcal{I}$ consists of the two preparation events P_A, P_B, and the initial beam splitting event S_i, while $\mathcal{F}$ consists of the recombination event S_f and the measurement events T_A and T_B.

The corresponding circuit diagram is obtained from the above one by promoting each event of interaction to a gate, and the propagation of each particle to a channel. This leads to the following circuit diagram (see Figure 14):

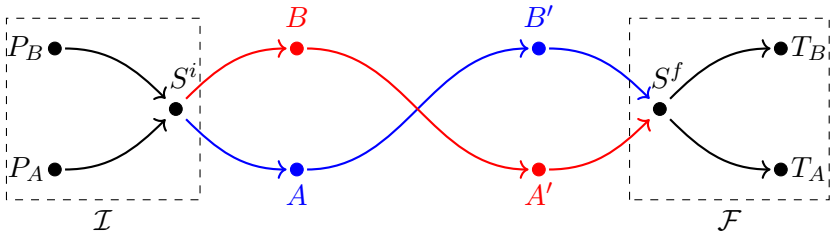


Figure 14: Circuit diagram of the 4-event implementation of the quantum switch. The internal structures of the composite gates $\mathcal{I}$ and $\mathcal{F}$ are explicitly depicted.

Its structure is in one-to-one correspondence with the spacetime diagram for the 4-event process, where the preparation and measurement spacetime events $\mathcal{I}$ and $\mathcal{F}$ have been split into three sub-gates each, for clarity.

The operations on each of the gates are given as follows. The preparation gate P_A maps from the input Hilbert space P_{A_I} to the output Hilbert space P_{A_O}, and analogously for gate P_B. The input spaces are trivial, $\dim P_{A_I} = \dim P_{B_I} = 1$, while each output space is spanned by vectors $|0\rangle$, $|1\rangle$ and $|v\rangle$. Here, $|0\rangle$ and $|1\rangle$ represent the two orthogonal qubit states (say, vertical and horizontal polarisations along certain axis in 3D space), while $|v\rangle$ is the vacuum state, representing the absence of particles in the corresponding arm of the interferometer. The operations performed at these gates, $P_A = |\Psi\rangle$ and $P_B = |v\rangle$, specify the initial conditions for the rest of the circuit diagram, and are described by the Choi-Jamiołkowski (CJ) states as

$$|P_A^*\rangle\!\rangle^{P_{A_I}P_{A_O}} = |\Psi^*\rangle^{P_{A_O}}\,, \qquad |P_B^*\rangle\!\rangle^{P_{B_I}P_{B_O}} = |v\rangle^{P_{B_O}}\,. \tag{35}$$

Here, $*$ denotes the complex conjugation.

Analogously, the target gates T_A and T_B facilitate the final measurement outcomes of the circuit diagram. The input spaces T_{A_I} and T_{B_I} are three-dimensional, spanned over the two qubit states and the vacuum, while the output spaces are one-dimensional. The operations performed at these gates, $T_\alpha = \langle\alpha|$ and $T_\beta = \langle\beta|$, read out the measurement results $\alpha, \beta \in \{0, 1, v\}$. The corresponding CJ states are given as

$$|T_\alpha^*\rangle\!\rangle^{T_{A_I}T_{A_O}} = |\alpha\rangle^{T_{A_I}}\,, \qquad |T_\beta^*\rangle\!\rangle^{T_{B_I}T_{B_O}} = |\beta\rangle^{T_{B_I}}\,. \tag{36}$$

The gates A, A', B and B' perform the unitaries U and V. The input and output spaces A_I and A_O of the Alice's gate A are both spanned by vectors $|0\rangle$, $|1\rangle$ and $|v\rangle$, and analogously for the input and output spaces of the remaining three gates. Assuming that in her (spatially) local laboratory Alice performs the unitary U on the particle's internal degree of freedom, the induced operation between the three-dimensional spaces A_I and A_O that include the vacuum states is given by

$$\tilde{U}^{A_OA_I} = U^{A_OA_I}P_{01}^{A_IA_I} + I^{A_OA_I}P_v^{A_IA_I}\,, \tag{37}$$

where $P_{01}^{A_IA_I} = |0\rangle^{A_I}\langle 0|^{A_I} + |1\rangle^{A_I}\langle 1|^{A_I}$, $P_v^{A_IA_I} = |v\rangle^{A_I}\langle v|^{A_I}$, and $I^{A_OA_I}$ represents the identity map between the Hilbert spaces A_O and A_I. The analogous construction also holds for the gate A', so the respective CJ states for the gates A and A' are then given by:

$$\begin{aligned} |\tilde{U}^*\rangle\!\rangle^{A_IA_O} &= \left[I^{A_IA_I}\otimes(\tilde{U}^*)^{A_OA_I}\right]|\mathbb{1}\rangle\!\rangle^{A_IA_I}\,,\\ |\tilde{U}^*\rangle\!\rangle^{A'_IA'_O} &= \left[I^{A'_IA'_I}\otimes(\tilde{U}^*)^{A'_OA'_I}\right]|\mathbb{1}\rangle\!\rangle^{A'_IA'_I}\,. \end{aligned} \tag{38}$$

Here, the "transport vector" is given by (for details of the process matrix formalism for the case of three-dimensional spaces — qutrits, see Appendix B):

$$|\mathbb{1}\rangle\!\rangle = |0\rangle|0\rangle + |1\rangle|1\rangle + |v\rangle|v\rangle\,. \tag{39}$$

Bob performs V in his (spatially) local laboratory, and therefore the CJ states for the gates B and B' are given as:

$$\begin{aligned} |\tilde{V}^*\rangle\!\rangle^{B_IB_O} &= \left[I^{B_IB_I}\otimes(\tilde{V}^*)^{B_OB_I}\right]|\mathbb{1}\rangle\!\rangle^{B_IB_I}\,,\\ |\tilde{V}^*\rangle\!\rangle^{B'_IB'_O} &= \left[I^{B'_IB'_I}\otimes(\tilde{V}^*)^{B'_OB'_I}\right]|\mathbb{1}\rangle\!\rangle^{B'_IB'_I}\,. \end{aligned} \tag{40}$$

The gates S^i and S^f act as beam splitters, i.e., they both perform the same Hadamard operation H, given as follows. The beam splitter input and output spaces consist of the Alice's and Bob's factor spaces. For the case of the Alice's input space, we have $S_{A_I} = \mathrm{span}\{|0\rangle^{S_{A_I}}, |1\rangle^{S_{A_I}}, |v\rangle^{S_{A_I}}\}$, and analogously for the output space, as well as for Bob's factor spaces. The overall input and output beam splitter spaces are therefore defined as $S_I = S_{(AB)_I} = S_{A_I}\otimes S_{B_I}$ and $S_O = S_{(AB)_O} = S_{A_O}\otimes S_{B_O}$. Finally, the unitary matrix associated to gate S representing the action of the balanced Hadamard beam splitter is given by:

$$\begin{aligned} H^{S_OS_I} = \;& \frac{1}{\sqrt{2}}\left(|0\rangle^{S_{A_O}}|v\rangle^{S_{B_O}} + |v\rangle^{S_{A_O}}|0\rangle^{S_{B_O}}\right)\langle 0|^{S_{A_I}}\langle v|^{S_{B_I}}\\ &+\frac{1}{\sqrt{2}}\left(|1\rangle^{S_{A_O}}|v\rangle^{S_{B_O}} + |v\rangle^{S_{A_O}}|1\rangle^{S_{B_O}}\right)\langle 1|^{S_{A_I}}\langle v|^{S_{B_I}}\\ &+\frac{1}{\sqrt{2}}\left(|0\rangle^{S_{A_O}}|v\rangle^{S_{B_O}} - |v\rangle^{S_{A_O}}|0\rangle^{S_{B_O}}\right)\langle v|^{S_{A_I}}\langle 0|^{S_{B_I}}\\ &+\frac{1}{\sqrt{2}}\left(|1\rangle^{S_{A_O}}|v\rangle^{S_{B_O}} - |v\rangle^{S_{A_O}}|1\rangle^{S_{B_O}}\right)\langle v|^{S_{A_I}}\langle 1|^{S_{B_I}}\,. \end{aligned} \tag{41}$$

The beam splitter acts such that the system coming from the Alice's side comes into an equal superposition of the two output spatial modes coming to Alice and Bob, with zero relative phase, while the system coming from the Bob's side (blue line) comes into an equal superposition of the two output

spatial modes with relative phase π. Thus, in the output space the correlation between the Alice's and Bob's vacuum state is the opposite as in the input case. The corresponding CJ state is then

$$|H^*\rangle\!\rangle^{S_I S_O} = \left[I^{S_I S_I} \otimes (H^*)^{S_O S_I}\right] |\mathbb{1}\rangle\!\rangle^{S_I S_I}\,, \tag{42}$$

where the transport vector $|\mathbb{1}\rangle\!\rangle$ for the beam splitter, when projected to a single-particle subspace, is given by

$$|\mathbb{1}\rangle\!\rangle = |0v\rangle|0v\rangle + |1v\rangle|1v\rangle + |v0\rangle|v0\rangle + |v1\rangle|v1\rangle\,. \tag{43}$$

Note that the full transport vector contains nine terms instead of the above four, but for the purpose of this paper, we do not need those five additional terms.

The process vector encodes the wires between the gates, and it is being constructed by taking the tensor product over appropriate transport vectors $|\mathbb{1}\rangle\!\rangle$ for Alice's and Bob's qutrits, see equations (26) and (27), such that each transport vector corresponds to one wire in the circuit diagram, connecting the output of the source gate to the input of the target gate. The process vector is thus given as:

$$\begin{aligned} |W_{\text{4-event}}\rangle\!\rangle \;=\; & \underbrace{|\mathbb{1}\rangle\!\rangle^{P_{A_O} S^i_{A_I}} |\mathbb{1}\rangle\!\rangle^{P_{B_O} S^i_{B_I}}}_{\text{initial}} \underbrace{|\mathbb{1}\rangle\!\rangle^{S^i_{A_O} A_I} |\mathbb{1}\rangle\!\rangle^{A_O B'_I} |\mathbb{1}\rangle\!\rangle^{B'_O S^f_{B_I}}}_{\text{blue}} \\ & \underbrace{|\mathbb{1}\rangle\!\rangle^{S^i_{B_O} B_I} |\mathbb{1}\rangle\!\rangle^{B_O A'_I} |\mathbb{1}\rangle\!\rangle^{A'_O S^f_{A_I}}}_{\text{red}} \underbrace{|\mathbb{1}\rangle\!\rangle^{S^f_{A_O} T_{A_I}} |\mathbb{1}\rangle\!\rangle^{S^f_{B_O} T_{B_I}}}_{\text{final}}\,. \end{aligned} \tag{44}$$

One can now evaluate the probability distribution

$$p(\alpha,\beta) = \left\| \mathcal{M}(\alpha,\beta) \right\|^2\,, \tag{45}$$

where the probability amplitude $\mathcal{M}(\alpha,\beta)$ is constructed by acting with the tensor product of all gate operations (35), (42), (38), (40), (42) and (36), on the process vector (44). Since each of the gate operations acts in its own part of the total Hilbert space, the order of application of these operations is immaterial, and we are free to choose the most convenient one.

To see what happens when the operations (35) of the preparation gates act on the process vector, let us evaluate the action of $|P^*_A\rangle\!\rangle^{P_{A_I} P_{A_O}}$ on $|\mathbb{1}\rangle\!\rangle^{P_{A_O} S^i_{A_I}}$:

$$\langle\!\langle P^*_A|^{P_{A_I} P_{A_O}} |\mathbb{1}\rangle\!\rangle^{P_{A_O} S^i_{A_I}} = \langle\Psi^*|^{P_{A_O}} \sum_{k=0}^{2} |k\rangle^{P_{A_O}} |k\rangle^{S^i_{A_I}} = \sum_{k=0}^{2} \Big(\langle\Psi|k\rangle\Big)^* |k\rangle^{S^i_{A_I}} = |\Psi\rangle^{S^i_{A_I}}\,. \tag{46}$$

An analogous calculation can be performed for $|P^*_B\rangle\!\rangle^{P_{B_I} P_{B_O}}$, so the action of both preparation operations (35) on the process vector (44) evaluates to:

$$\begin{aligned} \Big(\langle\!\langle P^*_A|^{P_{A_I} P_{A_O}} \otimes \langle\!\langle P^*_B|^{P_{B_I} P_{B_O}}\Big) |W_{\text{4-event}}\rangle\!\rangle = \\ |\Psi\rangle^{S^i_{A_I}} |v\rangle^{S^i_{B_I}} \underbrace{|\mathbb{1}\rangle\!\rangle^{S^i_{A_O} A_I} |\mathbb{1}\rangle\!\rangle^{A_O B'_I} |\mathbb{1}\rangle\!\rangle^{B'_O S^f_{B_I}}}_{\text{blue}} \\ \underbrace{|\mathbb{1}\rangle\!\rangle^{S^i_{B_O} B_I} |\mathbb{1}\rangle\!\rangle^{B_O A'_I} |\mathbb{1}\rangle\!\rangle^{A'_O S^f_{A_I}}}_{\text{red}} \underbrace{|\mathbb{1}\rangle\!\rangle^{S^f_{A_O} T_{A_I}} |\mathbb{1}\rangle\!\rangle^{S^f_{B_O} T_{B_I}}}_{\text{final}}\,. \end{aligned} \tag{47}$$

Next one acts with the initial Hadamard operation (42) on this process vector, transforming it into

$$\begin{aligned} \Big(\langle\!\langle P^*_A|^{P_{A_I} P_{A_O}} \otimes \langle\!\langle P^*_B|^{P_{B_I} P_{B_O}} \otimes \langle\!\langle S^*|^{S^i_{(AB)_I} S^i_{(AB)_O}}\Big) |W_{\text{4-event}}\rangle\!\rangle = \\ \frac{1}{\sqrt{2}} \Big(|\Psi\rangle^{A_I} |v\rangle^{B_I} + |v\rangle^{A_I} |\Psi\rangle^{B_I}\Big) \underbrace{|\mathbb{1}\rangle\!\rangle^{A_O B'_I} |\mathbb{1}\rangle\!\rangle^{B'_O S^f_{B_I}}}_{\text{blue}} \\ \underbrace{|\mathbb{1}\rangle\!\rangle^{B_O A'_I} |\mathbb{1}\rangle\!\rangle^{A'_O S^f_{A_I}}}_{\text{red}} \underbrace{|\mathbb{1}\rangle\!\rangle^{S^f_{A_O} T_{A_I}} |\mathbb{1}\rangle\!\rangle^{S^f_{B_O} T_{B_I}}}_{\text{final}} \equiv |W_{QS_4}\rangle\!\rangle\,. \end{aligned} \tag{48}$$

The resulting process vector is the outcome of the action of the composite gate $\mathcal{I}$ on (44), before the actions of Alice and Bob (note that often in the literature this is taken as the initial process vector in the analysis):

$$|W_{QS_4}\rangle\!\rangle = \frac{1}{\sqrt{2}}\left(|\Psi\rangle^{A_I}|v\rangle^{B_I} + |v\rangle^{A_I}|\Psi\rangle^{B_I}\right)|\mathbb{1}\rangle\!\rangle^{A_O B'_I}|\mathbb{1}\rangle\!\rangle^{B_O A'_I}|\mathbb{1}\rangle\!\rangle^{(A'_O B'_O)S_{(AB)_I}}|\mathbb{1}\rangle\!\rangle^{S_{(AB)_O}T_{(AB)_I}}. \quad (49)$$

Continuing the computation, the action of the remaining gate operations (38), (40), (36) and (42) on the process vector (49) gives us the probability amplitude,

$$\mathcal{M}(\alpha,\beta) \equiv$$
$$\left(\langle\!\langle\tilde{U}^*|^{A_I A_O} \otimes \langle\!\langle\tilde{U}^*|^{A'_I A'_O} \otimes \langle\!\langle\tilde{V}^*|^{B_I B_O} \otimes \langle\!\langle\tilde{V}^*|^{B'_I B'_O} \otimes \langle\!\langle H^*|^{S_I S_O} \otimes \langle\!\langle T_\alpha^*|^{T_{A_I} T_{A_O}} \otimes \langle\!\langle T_\beta^*|^{T_{B_I} T_{B_O}}\right)|W_{QS_4}\rangle\!\rangle . \quad (50)$$

Let us now calculate the action of $\langle\!\langle\tilde{U}^*|^{A_I A_O}$ on (49):

$$\langle\!\langle\tilde{U}^*|^{A_I A_O}|W_{QS_4}\rangle\!\rangle = \langle\!\langle\mathbb{1}|^{A_I A_I}\left[I^{A_I A_I} \otimes (\tilde{U}^T)^{A_O A_I}\right]|W_{QS_4}\rangle\!\rangle . \quad (51)$$

Looking at the structure of the process vector, one sees that the resulting new process vector will have the form

$$\langle\!\langle\tilde{U}^*|^{A_I A_O}|W_{QS_4}\rangle\!\rangle = \frac{1}{\sqrt{2}}\left(|X\rangle^{B'_I}|v\rangle^{B_I} + |Y\rangle^{B'_I}|\Psi\rangle^{B_I}\right)|\mathbb{1}\rangle\!\rangle^{B_O A'_I}|\mathbb{1}\rangle\!\rangle^{(A'_O B'_O)S_{(AB)_I}}|\mathbb{1}\rangle\!\rangle^{S_{(AB)_O}T_{(AB)_I}}, \quad (52)$$

where $|X\rangle^{B'_I}$ and $|Y\rangle^{B'_I}$ are shorthands for the expressions

$$|X\rangle^{B'_I} \equiv \langle\!\langle\mathbb{1}|^{A_I A_I}\left[I^{A_I A_I} \otimes (\tilde{U}^T)^{A_O A_I}\right]|\Psi\rangle^{A_I}|\mathbb{1}\rangle\!\rangle^{A_O B'_I} \quad (53)$$

and

$$|Y\rangle^{B'_I} \equiv \langle\!\langle\mathbb{1}|^{A_I A_I}\left[I^{A_I A_I} \otimes (\tilde{U}^T)^{A_O A_I}\right]|v\rangle^{A_I}|\mathbb{1}\rangle\!\rangle^{A_O B'_I}, \quad (54)$$

which need to be evaluated. The explicit computation of the first expression goes as follows:

$$\begin{aligned}
|X\rangle^{B'_I} &= \langle\!\langle\mathbb{1}|^{A_I A_I}\left[I^{A_I A_I} \otimes (\tilde{U}^T)^{A_O A_I}\right]|\Psi\rangle^{A_I}|\mathbb{1}\rangle\!\rangle^{A_O B'_I} \\
&= \sum_{k=0}^{2}\langle k|^{A_I}\langle k|^{A_I}\left[I^{A_I A_I} \otimes (\tilde{U}^T)^{A_O A_I}\right]|\Psi\rangle^{A_I}\sum_{m=0}^{2}|m\rangle^{A_O}|m\rangle^{B'_I} \\
&= \sum_{m=0}^{2}\left[\sum_{k=0}^{2}\left(\langle k|^{A_I}I^{A_I A_I}|\Psi\rangle^{A_I}\right)\left(\langle k|^{A_I}(\tilde{U}^T)^{A_O A_I}|m\rangle^{A_O}\right)\right]|m\rangle^{B'_I} \\
&= \sum_{m=0}^{2}\left[\sum_{k=0}^{2}\langle k|\Psi\rangle^{A_I}\,\langle m|^{A_O}\tilde{U}^{A_O A_I}|k\rangle^{A_I}\right]|m\rangle^{B'_I} \\
&= \sum_{m=0}^{2}\left[\langle m|^{A_O}\tilde{U}^{A_O A_I}|\Psi\rangle^{A_I}\right]|m\rangle^{B'_I}.
\end{aligned} \quad (55)$$

Using (37), the coefficient in the brackets can be evaluated as

$$\langle m|^{A_O}\tilde{U}^{A_O A_I}|\Psi\rangle^{A_I} = \langle m|^{A_O}\left(U^{A_O A_I}P_{01}^{A_I A_I} + I^{A_O A_I}P_v^{A_I A_I}\right)|\Psi\rangle^{A_I} = \langle m|U|\Psi\rangle, \quad (56)$$

since $P_{01}^{A_I A_I}|\Psi\rangle^{A_I} = |\Psi\rangle^{A_I}$ and $P_v^{A_I A_I}|\Psi\rangle^{A_I} = 0$. Thus, we have

$$|X\rangle^{B'_I} = \sum_{m=0}^{2}\langle m|U|\Psi\rangle\,|m\rangle^{B'_I} = U|\Psi\rangle^{B'_I} \equiv |U\Psi\rangle^{B'_I}. \quad (57)$$

The computation of $|Y\rangle^{B'_I}$ proceeds in an analogous way to (55), and the result is

$$|Y\rangle^{B'_I} = \sum_{m=0}^{2} \left[\langle m|^{A_O} \tilde{U}^{A_O A_I} |v\rangle^{A_I} \right] |m\rangle^{B'_I} . \tag{58}$$

Again, using (37), the coefficient in the brackets can be evaluated as

$$\langle m|^{A_O} \tilde{U}^{A_O A_I} |v\rangle^{A_I} = \langle m|^{A_O} \left(U^{A_O A_I} P_{01}^{A_I A_I} + I^{A_O A_I} P_v^{A_I A_I} \right) |v\rangle^{A_I} = \langle m|v\rangle = \delta_{mv} , \tag{59}$$

since $P_{01}^{A_I A_I} |v\rangle^{A_I} = 0$ and $P_v^{A_I A_I} |v\rangle^{A_I} = |v\rangle^{A_I}$. Thus, we have

$$|Y\rangle^{B'_I} = \sum_{m=0}^{2} \delta_{mv} \, |m\rangle^{B'_I} = |v\rangle^{B'_I} . \tag{60}$$

Finally, substituting (57) and (60) back into (52), we obtain:

$$\langle\!\langle \tilde{U}^* |^{A_I A_O} |W_{QS_4}\rangle\!\rangle = \frac{1}{\sqrt{2}} \left(|U\Psi\rangle^{B'_I} |v\rangle^{B_I} + |v\rangle^{B'_I} |\Psi\rangle^{B_I} \right) |\mathbb{1}\rangle\!\rangle^{B_O A'_I} |\mathbb{1}\rangle\!\rangle^{(A'_O B'_O) S_{(AB)_I}} |\mathbb{1}\rangle\!\rangle^{S_{(AB)_O} T_{(AB)_I}} . \tag{61}$$

One should note, comparing (61) with (49), that the action of the gate A operation onto the process vector effectively performs the following transformation,

$$|\Psi\rangle^{A_I} \to |U\Psi\rangle^{A_O} \to |U\Psi\rangle^{B'_I} , \qquad |v\rangle^{A_I} \to |v\rangle^{A_O} \to |v\rangle^{B'_I} , \tag{62}$$

where the transport vector $|\mathbb{1}\rangle\!\rangle^{A_O B'_I}$ has been utilised for "transporting" the state from the output A_O of gate A to the input B'_I of the gate B', in line with the spacetime diagram. This scheme repeats itself with the action of all remaining gate operations on (61). In particular, the subsequent action of the gate B operation gives:

$$\left(\langle\!\langle \tilde{V}^* |^{B_I B_O} \otimes \langle\!\langle \tilde{U}^* |^{A_I A_O} \right) |W_{QS_4}\rangle\!\rangle =$$

$$\frac{1}{\sqrt{2}} \left(|U\Psi\rangle^{B'_I} |v\rangle^{A'_I} + |v\rangle^{B'_I} |V\Psi\rangle^{A'_I} \right) |\mathbb{1}\rangle\!\rangle^{(A'_O B'_O) S_{(AB)_I}} |\mathbb{1}\rangle\!\rangle^{S_{(AB)_O} T_{(AB)_I}} , \tag{63}$$

which can also be verified with an explicit calculation similar to the above. Continuing on, the operations at the gates A' and B' give:

$$\begin{aligned} &\left(\langle\!\langle \tilde{V}^* |^{B'_I B'_O} \otimes \langle\!\langle \tilde{U}^* |^{A'_I A'_O} \otimes \langle\!\langle \tilde{V}^* |^{B_I B_O} \otimes \langle\!\langle \tilde{U}^* |^{A_I A_O} \right) |W_{QS_4}\rangle\!\rangle = \\ &\qquad \frac{1}{\sqrt{2}} \left(|VU\Psi\rangle^{S_{B_I}} |v\rangle^{S_{A_I}} + |v\rangle^{S_{B_I}} |UV\Psi\rangle^{S_{A_I}} \right) |\mathbb{1}\rangle\!\rangle^{S_{(AB)_O} T_{(AB)_I}} . \end{aligned} \tag{64}$$

Next, the action of the beam splitter at the gate S_f gives

$$\begin{aligned} &\left(\langle\!\langle \tilde{H}^* |^{S_I S_O} \otimes \langle\!\langle \tilde{V}^* |^{B'_I B'_O} \otimes \langle\!\langle \tilde{U}^* |^{A'_I A'_O} \otimes \langle\!\langle \tilde{V}^* |^{B_I B_O} \otimes \langle\!\langle \tilde{U}^* |^{A_I A_O} \right) |W_{QS_4}\rangle\!\rangle = \\ &\qquad \frac{1}{2} \sum_{i=0}^{1} \left(\langle i| \{ U , V \} |\Psi\rangle \, |i\rangle^{T_{A_I}} |v\rangle^{T_{B_I}} + \langle i| [U , V] |\Psi\rangle \, |v\rangle^{T_{A_I}} |i\rangle^{T_{B_I}} \right) . \end{aligned} \tag{65}$$

Finally, the action of the operations of the target gates T_A and T_B gives us the probability amplitude as a function of the measurement outcomes α and β,

$$\mathcal{M}(\alpha, \beta) \equiv$$

$$\left(\langle\!\langle \tilde{U}^* |^{A_I A_O} \otimes \langle\!\langle \tilde{U}^* |^{A'_I A'_O} \otimes \langle\!\langle \tilde{V}^* |^{B_I B_O} \otimes \langle\!\langle \tilde{V}^* |^{B'_I B'_O} \otimes \langle\!\langle H^* |^{S_I S_O} \otimes \langle\!\langle T_\alpha^* |^{T_{A_I} T_{A_O}} \otimes \langle\!\langle T_\beta^* |^{T_{B_I} T_{B_O}} \right) |W_{QS_4}\rangle\!\rangle$$

$$= \frac{1}{2}\Big[\delta_{\beta v}\langle\alpha|\{U\,,V\}|\Psi\rangle + \delta_{\alpha v}\langle\beta|[U\,,V]|\Psi\rangle\Big]\,. \tag{66}$$

At this point we can employ (45) and calculate the probability distribution,

$$p(\alpha,\beta) = \frac{1}{4}\Big[\delta_{\beta v}\Big|\langle\alpha|\{U\,,V\}|\Psi\rangle\Big|^2 + \delta_{\alpha v}\Big|\langle\beta|[U\,,V]|\Psi\rangle\Big|^2\Big]\,, \tag{67}$$

where we have used the fact that the vacuum state $|v\rangle$ is orthogonal to both $\{U\,,V\}|\Psi\rangle$ and $[U\,,V]|\Psi\rangle$. In particular, we see that for $i\in\{0,1\}$ we have

$$p(i,v) = \frac{1}{4}\Big|\langle i|\{U\,,V\}|\Psi\rangle\Big|^2\,, \qquad p(v,i) = \frac{1}{4}\Big|\langle i|[U\,,V]|\Psi\rangle\Big|^2\,, \tag{68}$$

while all other choices of α and β give $p(\alpha,\beta)=0$. The total probability that Alice will detect a particle is given by the marginal

$$p_A = \sum_{i=1}^{2} p(i,v) = \frac{1}{2}\Big(1+\mathrm{Re}\,\langle\Psi|U^\dagger V^\dagger U V|\Psi\rangle\Big)\,, \tag{69}$$

while the corresponding total probability that Bob will detect a particle is

$$p_B = \sum_{i=1}^{2} p(v,i) = \frac{1}{2}\Big(1-\mathrm{Re}\,\langle\Psi|U^\dagger V^\dagger U V|\Psi\rangle\Big)\,. \tag{70}$$

As a final point, we can verify that the probability distribution is correctly normalised. Using the fact that the only nonzero values for the probability are given in (68), we have

$$p_{\text{total}} = \sum_{\alpha=0}^{2}\sum_{\beta=0}^{2} p(\alpha,\beta) = \underbrace{p(0,v)+p(1,v)}_{p_A} + \underbrace{p(v,0)+p(v,1)}_{p_B} = 1\,, \tag{71}$$

as expected.

Instead of recombining the particle on the second beam splitter, one can consider the case in which the final gate $\mathcal{F}$ consists only of local measurements performed onto a particle in the Alice's and Bob's paths. In this case, the final gate is equivalent to two target gates T_A and T_B. Given that all the processes considered are pure, the corresponding process vector for the 4-event quantum switch implementation is given as (in order to compare the current with the previous works, we present the process that starts after $\mathcal{I}$, as was done in, say, [9]):

$$|W_{QS_4}\rangle\!\rangle = \frac{1}{\sqrt{2}}\Big(|\Psi\rangle^{A_I}|v\rangle^{B_I} + |v\rangle^{A_I}|\Psi\rangle^{B_I}\Big)\,|\mathbb{1}\rangle\!\rangle^{A_O B'_I}|\mathbb{1}\rangle\!\rangle^{B_O A'_I}|\mathbb{1}\rangle\!\rangle^{A'_O T_{A_I}}|\mathbb{1}\rangle\!\rangle^{B'_O T_{B_I}}\,. \tag{72}$$

D 3-event process vector

The detailed spacetime diagram for the 3-event quantum switch is given below (see Figure 15).

The process vector for this case is obtained from (44) by identifying the spacetime positions of the gates B and B', yet keeping the corresponding Hilbert spaces in the mathematical description (i.e., keeping the dimensionality of the problem). Thus, the corresponding circuit is *identical* to the 4-event circuit, and the process vector has the *identical* mathematical form as in the case of four gates. In order to emphasise the physical difference between the two cases, instead of $B_{I/O}$ and $B'_{I/O}$, we write B_{I_1/O_1} and B_{I_2/O_2}, respectively:

$$\begin{aligned} |W_{\text{3-event}}\rangle\!\rangle \;=\; & \underbrace{|\mathbb{1}\rangle\!\rangle^{P_{A_O}S^i_{A_I}}|\mathbb{1}\rangle\!\rangle^{P_{B_O}S^i_{B_I}}}_{\text{initial}}\underbrace{|\mathbb{1}\rangle\!\rangle^{S^i_{A_O}A_I}|\mathbb{1}\rangle\!\rangle^{A_O B_{I_2}}|\mathbb{1}\rangle\!\rangle^{B_{O_2}S^f_{B_I}}}_{\text{blue}} \\ & \underbrace{|\mathbb{1}\rangle\!\rangle^{S^i_{B_O}B_{I_1}}|\mathbb{1}\rangle\!\rangle^{B_{O_1}A'_I}|\mathbb{1}\rangle\!\rangle^{A'_O S^f_{A_I}}}_{\text{red}}\underbrace{|\mathbb{1}\rangle\!\rangle^{S^f_{A_O}T_{A_I}}|\mathbb{1}\rangle\!\rangle^{S^f_{B_O}T_{B_I}}}_{\text{final}}\,. \end{aligned} \tag{73}$$

The final probability distribution is identical to the one for the 4-event process, given by (67).

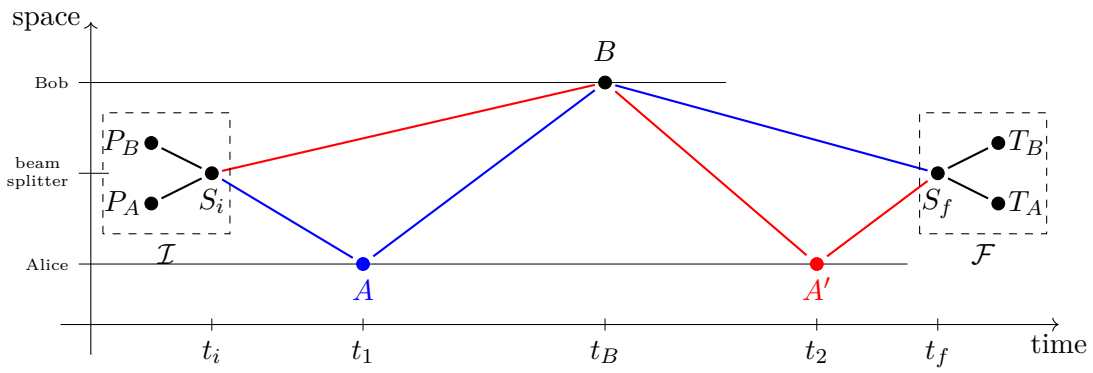


Figure 15: Spacetime diagram of the 3-event implementation of the quantum switch. The internal structures of the composite gates $\mathcal{I}$ and $\mathcal{F}$ are explicitly depicted.

E 2-event

In this Appendix we present process vectors for the two gravitational switches discussed in the main text. First, the process vector of the gravitational switch without recombination [15], is given by (since the "control" is now played by gravity, it is thus denoted as G, instead of C):

$$|W_{QS_2}\rangle\!\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle^{G}|\Psi\rangle^{A_I}|\mathbb{1}\rangle\!\rangle^{A_O B_I}|\mathbb{1}\rangle\!\rangle^{B_O T_{B_I}} + |1\rangle^{G}|\Psi\rangle^{B_I}|\mathbb{1}\rangle\!\rangle^{B_O A_I}|\mathbb{1}\rangle\!\rangle^{A_O T_{A_I}}\right)|\mathbb{1}\rangle\!\rangle^{G T_{G_I}}. \tag{74}$$

It is then straightforward to obtain the process vector for the case of recombining only the gravity on the final beam splitter S_f (a part of a bigger final gate $\mathcal{F}$), obtaining the analogue of (49), while the particle is not being recombined. Note that in this case the introduction of the vacuum state is not necessary, as in each branch of superposition all gates are acting upon a particle.

In contrast to the above case, the process vector describing the gravitational 2-event quantum switch with the recombination of both gravity and the particle is given as follows:

$$\begin{aligned}|W^{(r)}_{QS_2}\rangle\!\rangle = \frac{1}{\sqrt{2}}&\left(|0\rangle^{G}|\Psi\rangle^{A_I}|\mathbb{1}\rangle\!\rangle^{A_O B_I}|\mathbb{1}\rangle\!\rangle^{B_O S_{P_I}} + |1\rangle^{G}|\Psi\rangle^{B_I}|\mathbb{1}\rangle\!\rangle^{B_O A_I}|\mathbb{1}\rangle\!\rangle^{A_O S_{P_I}}\right)\\ &\otimes|\mathbb{1}\rangle\!\rangle^{G S_{G_I}}|\mathbb{1}\rangle\!\rangle^{(S_{G_O}S_{P_O})(T_{G_I}T_{P_I})}.\end{aligned} \tag{75}$$

Here, P stands for "the particle" (whose corresponding input space S_{P_I} is isomorphic to the tensor product of Alice's and Bob's output spaces, $S_{P_I} \simeq A_O \otimes B_O$), and S stands for a "beam splitter" (whose corresponding input space is $S_I = S_{G_I} \otimes S_{P_I}$, and analogously for the output space).

While defining the spaces $S_{G_{I/O}}$, $S_{P_{I/O}}$, T_{G_I}, T_{P_I}, and the vector $|\mathbb{1}\rangle\!\rangle^{(S_{G_O}S_{P_O})(T_{G_I}T_{P_O})}$ is straightforward, it is not so for the "final" recombination vector $|U^*_{BS}\rangle\!\rangle^{(S_{G_I}S_{P_I})(S_{G_O}S_{P_O})}$. Namely, note that in our gravitational switch *all* degrees of freedom, both gravitational and matter, are recombined by U_{BS} such that, by acting on the beam splitter input entangled state,

$$|\Psi_i\rangle^{S_{G_I}S_{P_I}} = \frac{1}{\sqrt{2}}\left(|0\rangle^{S_{G_I}} \otimes \left[UV|\Psi\rangle^{S_{P_I}}\right] + |1\rangle^{S_{G_I}} \otimes \left[VU|\Psi\rangle^{S_{P_I}}\right]\right), \tag{76}$$

the overall output state is a product one, of the form

$$|\Psi_o\rangle^{S_{G_O}S_{P_O}} = |\Gamma\rangle^{S_{G_O}} \otimes (\alpha UV + \beta VU)\,|\Psi\rangle^{S_{P_O}}, \tag{77}$$

where $|\Gamma\rangle^{S_{G_O}}$ is some (not necessarily classical) state of gravity. The above evolution is, at least in principle, allowed by the quantum laws, which is all that we need to know regarding the action of U_{BS} at this point. Its action on the rest of the overall Hilbert space is, for the purpose of our argument, irrelevant, and can thus be chosen arbitrarily.

Finally, we would like to note that the same type of the 4-, or 3-event quantum switches in classical spacetimes can also be defined, resulting in the same type of the output state as (77), with the gravity degree of freedom being replaced by any additional matter degree of freedom that plays the role of the control.

F Various implementations of the gravitational switch

In this Appendix we present a few representative additional constructions of the gravitational quantum switch. First, we start with a 2-event switch for which the requirement (i) from Subsection 4.3 is not satisfied. It is obvious from the diagram on the left that each of the photon's superposed trajectories fail to meet at the boundary region, violating requirement (i), see the left diagram of Figure 16. Next, we proceed with the 2-event implementation for which requirement (i) is satisfied, but the requirement (ii) is not, since the superposed trajectories fail to recombine. This is depicted on the right diagram of Figure 16.

Finally, we present a 4-event implementation for which *both* requirements (i) and (ii) are satisfied (see Figure 17).

Of course, other variations are possible as well.

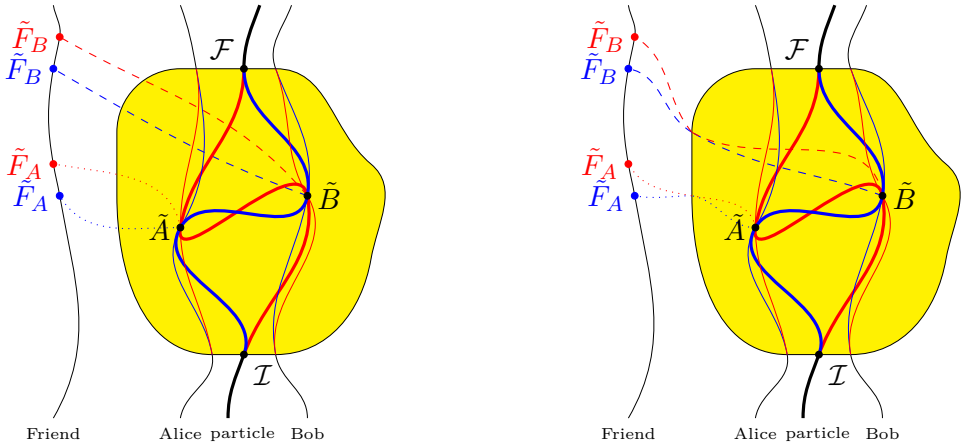


Figure 16: The spacetime diagrams of a 2-event gravitational switch implementations, with Friend's measurements, which fail to distinguish them from the optical implementation of the quantum switch.

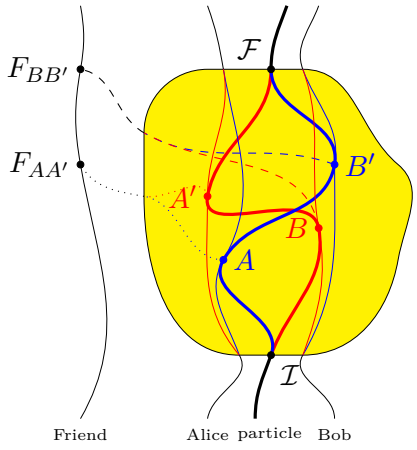


Figure 17: The spacetime diagram of a 4-event gravitational switch implementation, with Friend's measurement, which fails to distinguish it from the 2-event implementation of the gravitational switch.

Quantum Information Processing (2019) 18:176
https://doi.org/10.1007/s11128-019-2293-7

Creating photonic GHZ and W states via quantum walk

Le Ju[1] · Ming Yang[1] · Nikola Paunković[2,3,4] · Wen-Jing Chu[1] · Zhuo-Liang Cao[1,5]

Received: 5 November 2018 / Accepted: 17 April 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
We propose schemes for preparing W and GHZ photonic states entangled in polarization via quantum walks, where the trajectory exchanging of identical photons plus postselection induces the coupling between their polarization degrees of freedom. Being different from the polarization beam splitter-based approach, the indistinguishability of identical photons provides an alternative for coupling photons, which is more phase stable and may induce richer actions than the traditional polarization beam splitter on input photons. Our schemes demonstrate that this new coupling mechanism for photons is originated from the fact that trajectory-exchanging operation plus postselection can extract entangled states from the initially product states, and this new coupling mechanism may find considerable potential applications in both quantum communication and quantum computation.

Keywords Generation of entangled states · W state · GHZ state · Quantum walk

1 Introduction

Quantum entanglement plays an important role in quantum computation and quantum communication [1]. The most common are bipartite entangled states, which have

✉ Ming Yang
mingyang@ahu.edu.cn

Nikola Paunković
npaunkovic@gmail.com

1 School of Physics and Materials Science, Anhui University, Hefei 230601, China

2 Instituto de Telecomunicações, Av. Rovisco Pais, 1049-001 Lisboa, Portugal

3 Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal

4 CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal

5 School of Electronic and Information Engineering, Hefei Normal University, Hefei 230601, China

Published online: 25 April 2019

been widely utilized in many quantum information processes. Regarding multipartite entangled states, it was found that they have stronger computational power than the bipartite counterparts [2]. Because the structure of multipartite entangled states is more complex than that of bipartite counterparts, their preparation is more difficult. In 2000, Dür *et. al.* [3] proved that there are two inequivalent three-qubit entangled states that exhibit genuine tripartite entanglement—W state and GHZ state, and they cannot be converted into each other by local operations and classical communication (LOCC). The component particles of W state are in identical subsystem partial states, and the entanglement between every pair of component particles reaches the upper bound achievable. Even if one component particle of the W state is lost, the remaining component particles are still entangled [4,5], so the W state is resistant to particle loss. Furthermore, W state is considered as an entanglement resource for the optimal quantum cryptography [6,7], quantum teleportation [8–10], and quantum dense coding [10,11]. As a maximally entangled state, GHZ state is the best quantum channel for teleportation [12] and quantum key distribution [13], and it can achieve the Heisenberg limit in quantum metrology [14]. Many people worked on the relationship between GHZ state and the Bell theorem [15,16] and studied the violation of the Bell inequality experimentally [17].

How to effectively achieve the construction of multipartite entangled quantum states using the existing physical systems is an important question. Many schemes have been proposed to achieve the preparation of multipartite entangled states, which can be divided into two categories. One uses solid-state systems, such as ion traps [18,19], superconducting circuits [20–22] cavity quantum electrodynamics (cavity QED) [23–26], Ising spin chains [27,28] and so on. But the harsh experimental conditions of these schemes make them rather complicated. The other uses the linear optical systems, such as spontaneous parametric down-conversion (SPDC) [29,30], expansion schemes [31,32], fusion schemes [33–40], and so on. Recent experiment has achieved the production of 10-photon [41] and 18-qubit GHZ states [42].

The underlying physical mechanism for most of these currently existing entanglement generation schemes is the dynamical interaction between particles. But to let the elementary systems interact is not an easy task. Besides the interaction-based entanglement generation mechanism, particle indistinguishability can be regarded as an effective resource of entanglement too (see Chapter 9.2 of [43] and [44,45]). It was first utilized to test the GHZ [46] and EPR [47] correlations, and it was recently used to generate the entangled states for quantum information processing, such as Bell state [48,49], GHZ state [48,49], and W state [49] in a non-dynamical (interactionless) manner. Most of these generated states are entangled in the spatial degree of freedom, and the entangled states for the internal pseudospin degree of freedom are more useful than those entangled in the spatial degree of freedom. Although the photons are entangled in the polarization degree of freedom in the scheme of Ref. [48], the scheme cannot be used to generate polarization-entangled W state. So it is of great importance to design a new scheme to generate polarization-entangled W state via particle indistinguishability rather than a dynamical interaction between photons. Meanwhile, we found that quantum walk can be used to implement entanglement swapping for photonic entangled states; that is to say, Bell state measurement in entanglement swapping can be realized by using the indistinguishability of photons during optical quantum

walks [50], which indicates that it is possible to design a quantum walk-based scheme to generate polarization entangle W state via particle indistinguishability rather than a dynamical interaction between photons.

Quantum walk is a quantum generalization of classical random walk, which has attracted considerable attention after being put forward by Aharonov for the first time in 1993 [51]. Quantum walk has been proved to be feasible. Many works have been done on it [52–55], such as quantum-to-classical transition [56,57], quantum measurement [58–60], quantum computation [61–64], quantum teleportation [65], and entanglement swapping [50] etc. In an optical quantum walk, the photons with orthogonal polarization states will walk to different positions, respectively, and the trajectory-exchanging process plus postselection during optical quantum walks can be used to couple photons without the beam splitter of the standard coupling scheme [50]. In this paper, we propose a new scheme to generate polarization entangle W state via the trajectory-exchanging process plus postselection during optical quantum walks, which makes a full use of the indistinguishability. The results show that a slightly modified version of it can generate polarization GHZ state too, and the efficiency of our W state generation scheme is higher than that of Ref. [49].

In addition, the parameters of the beam splitter of the standard coupling schemes cannot be modified after being fabricated, which may limit the coupling mechanisms for photons. Because our entanglement generation scheme is designed without beam splitter, it will greatly enrich the coupling mechanism for photons.

2 Preparation of three-photon entangled states

To entangle three independent polarized photons, we have to input them into the setup depicted in Fig. 1 through different positions, and the input state of the three identical photons can be expressed as:

$$|\psi^{(0)}\rangle = a^{\dagger}_{H,1} a^{\dagger}_{H,0} a^{\dagger}_{H,-1} |0\rangle, \tag{1}$$

where $|0\rangle$ is the vacuum state, and φ and x in $a^{\dagger}_{\varphi,x}$ label the polarization state and the current position of the photon on the beam displacer (BD in Fig. 1), respectively. Thus, φ can be H and V, representing horizontal and vertical polarization states, respectively, and $x = -2, -1, 0, +1, +2$.

The setup depicted in Fig. 1 is based on the quantum walk of photons induced by the BDs plus the wave plates. In an optical quantum walk, the photons with orthogonal polarization states at the same position will arrive at opposite positions with respect to the initial one, which is realized by a polarization rotation operation U and a subsequent conditional position shift operation S. Different polarization rotation operations U can be implemented by half-wave plate (HWP) set at different angles δ:

$$HWP_{\delta} = \begin{pmatrix} \cos 2\delta & \sin 2\delta \\ \sin 2\delta & -\cos 2\delta \end{pmatrix}. \tag{2}$$

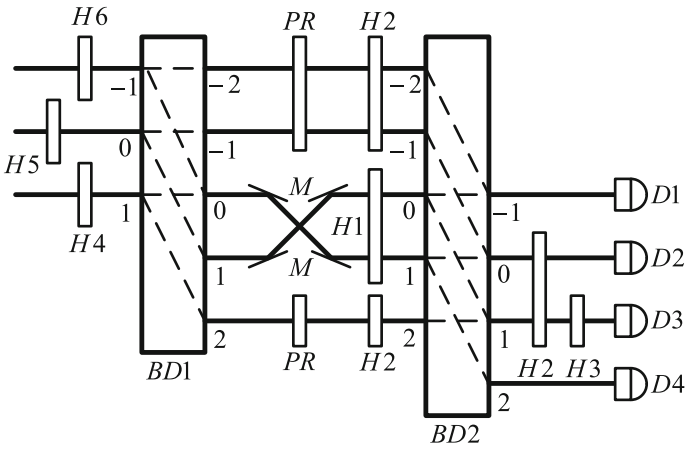


Fig. 1 The optical circuit for preparing polarization-entangled W state. BD indicates beam displacer, M denotes mirror, and PR indicates phase retarder, which is used for phase compensation. $Hi(i = 1, 2, \cdots, 6)$ are half-wave plate, which are set at different angles to implement different polarization rotation operations Ui in Eq. (3), respectively

Corresponding to the six HWPs in Fig. 1, the polarization rotation operations U are expressed as:

$$
\begin{aligned}
U1 &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, U2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, U3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\
U4 &= \begin{pmatrix} \sqrt{\frac{1}{\sqrt{3}}} & \sqrt{1-\frac{1}{\sqrt{3}}} \\ \sqrt{1-\frac{1}{\sqrt{3}}} & -\sqrt{\frac{1}{\sqrt{3}}} \end{pmatrix}, \\
U5 &= \begin{pmatrix} \sqrt{2-\frac{2}{\sqrt{3}}} & \sqrt{\frac{2}{\sqrt{3}}-1} \\ \sqrt{\frac{2}{\sqrt{3}}-1} & -\sqrt{2-\frac{2}{\sqrt{3}}} \end{pmatrix}, \\
U6 &= \begin{pmatrix} \sqrt{1-\frac{1}{\sqrt{3}}} & \sqrt{\frac{1}{\sqrt{3}}} \\ \sqrt{\frac{1}{\sqrt{3}}} & -\sqrt{1-\frac{1}{\sqrt{3}}} \end{pmatrix}.
\end{aligned} \tag{3}
$$

The conditional position shift operation S can be realized by a birefringent calcite beam displacer (BD), which displaces the horizontal polarization component (H) away from the vertical polarization component (V). The function of a BD can be expressed as:

$$S = \sum_x (|H, x+1\rangle\langle H, x| + |V, x-1\rangle\langle V, x|). \tag{4}$$

Another key operation in our scheme is the trajectory-exchanging operation whose function can be defined as follows. Photons with polarization states φ_1 and φ_2 are initially located at positions x and y, respectively, and the trajectory-exchanging oper-

ation acts as:

$$\begin{aligned} a^{\dagger}_{\varphi_1,x}|0\rangle &\rightarrow a^{\dagger}_{\varphi_1,y}|0\rangle, \\ a^{\dagger}_{\varphi_2,y}|0\rangle &\rightarrow a^{\dagger}_{\varphi_2,x}|0\rangle. \end{aligned} \tag{5}$$

Let's move to a detailed description of the scheme. As depicted in Fig. 1, the input photon at position 1 passes through $H4$, the input photon at position 0 through $H5$, and the input photon at position -1 through $H6$. After passing through corresponding HWPs, three independent photons enter the first beam displacer $BD1$. The output modes $+1$ and 0 of $BD1$ are exchanged via four mirrors, and enter the $+1$ and 0 modes of $BD2$, after passing through $H1$. Other output modes $(-2, -1, +2)$ of $BD1$ successively pass through phase retarder and $H2$ and then enter the corresponding modes of $BD2$. An $H2$ is added at the 0 output port of $BD2$, and $H2$ plus $H3$ are added at the $+1$ output port. At the output port of the setup, the three photons are in the following state:

$$\begin{aligned} |\psi^{(2)}\rangle = & \left[\sqrt{\frac{1}{\sqrt{3}}}a^{\dagger}_{H,1} + \sqrt{\frac{1}{2} - \frac{1}{2\sqrt{3}}}(a^{\dagger}_{H,2} - a^{\dagger}_{H,0})\right] \\ & \left[\sqrt{1 - \frac{1}{\sqrt{3}}}(-a^{\dagger}_{V,1} + a^{\dagger}_{V,-1}) + \sqrt{\frac{2}{\sqrt{3}} - 1}a^{\dagger}_{V,0}\right] \\ & \bigotimes \left[\sqrt{\frac{1}{2} - \frac{1}{2\sqrt{3}}}(a^{\dagger}_{H,2} + a^{\dagger}_{H,0}) + \sqrt{\frac{1}{\sqrt{3}}}a^{\dagger}_{H,-1}\right]|0\rangle \\ = & \sqrt{\frac{2}{\sqrt{3}} - 1}\left[\frac{1}{\sqrt{3}}(a^{\dagger}_{H,1}a^{\dagger}_{H,0}a^{\dagger}_{V,-1} + a^{\dagger}_{H,1}a^{\dagger}_{V,0}a^{\dagger}_{H,-1} + a^{\dagger}_{V,1}a^{\dagger}_{H,0}a^{\dagger}_{H,-1})|0\rangle\right] \\ & + |\phi\rangle_{\text{other}} \\ = & \sqrt{\frac{2}{\sqrt{3}} - 1}|W\rangle_3 + |\phi\rangle_{\text{other}}, \end{aligned} \tag{6}$$

where

$$\begin{aligned} |\phi\rangle_{\text{other}} = & \left(-\sqrt{\frac{2}{3\sqrt{3}} - \frac{1}{3}}a^{\dagger}_{H,1}a^{\dagger}_{V,1}a^{\dagger}_{H,2} - \sqrt{\frac{2}{3\sqrt{3}} - \frac{1}{3}}a^{\dagger}_{H,1}a^{\dagger}_{V,1}a^{\dagger}_{H,0}\right. \\ & - \sqrt{\frac{1}{3} - \frac{1}{3\sqrt{3}}}a^{\dagger}_{H,1}a^{\dagger}_{V,1}a^{\dagger}_{H,-1} \\ & + \sqrt{\frac{2}{3\sqrt{3}} - \frac{1}{3}}a^{\dagger}_{H,1}a^{\dagger}_{V,-1}a^{\dagger}_{H,2} + \sqrt{\frac{1}{3} - \frac{1}{3\sqrt{3}}}a^{\dagger}_{H,1}a^{\dagger}_{V,-1}a^{\dagger}_{H,-1} \\ & \left. + \sqrt{\frac{1}{2} - \frac{5}{6\sqrt{3}}}a^{\dagger}_{H,1}a^{\dagger}_{V,0}a^{\dagger}_{H,2}\right. \end{aligned}$$

$$
\begin{aligned}
&+\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,1}a^{\dagger}_{V,0}a^{\dagger}_{H,0}-\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,2}a^{\dagger}_{V,1}a^{\dagger}_{H,2}\\
&-\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,2}a^{\dagger}_{V,1}a^{\dagger}_{H,0}\\
&-\sqrt{\frac{2}{3\sqrt{3}}-\frac{1}{3}}a^{\dagger}_{H,2}a^{\dagger}_{V,1}a^{\dagger}_{H,-1}+\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,2}a^{\dagger}_{V,-1}a^{\dagger}_{H,2}\\
&+\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,2}a^{\dagger}_{V,-1}a^{\dagger}_{H,0}\\
&+\sqrt{\frac{2}{3\sqrt{3}}-\frac{1}{3}}a^{\dagger}_{H,2}a^{\dagger}_{V,-1}a^{\dagger}_{H,-1}+\sqrt{\frac{7}{6\sqrt{3}}-\frac{2}{3}}a^{\dagger}_{H,2}a^{\dagger}_{V,0}a^{\dagger}_{H,2}\\
&+\sqrt{\frac{7}{6\sqrt{3}}-\frac{2}{3}}a^{\dagger}_{H,2}a^{\dagger}_{V,0}a^{\dagger}_{H,0}\\
&+\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,2}a^{\dagger}_{V,0}a^{\dagger}_{H,-1}+\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,0}a^{\dagger}_{V,1}a^{\dagger}_{H,2}\\
&+\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,0}a^{\dagger}_{V,1}a^{\dagger}_{H,0}\\
&-\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,0}a^{\dagger}_{V,-1}a^{\dagger}_{H,2}-\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,0}a^{\dagger}_{V,-1}a^{\dagger}_{H,0}\\
&-\sqrt{\frac{2}{3\sqrt{3}}-\frac{1}{3}}a^{\dagger}_{H,0}a^{\dagger}_{V,-1}a^{\dagger}_{H,-1}\\
&-\sqrt{\frac{7}{6\sqrt{3}}-\frac{2}{3}}a^{\dagger}_{H,0}a^{\dagger}_{V,0}a^{\dagger}_{H,2}-\sqrt{\frac{7}{6\sqrt{3}}-\frac{2}{3}}a^{\dagger}_{H,0}a^{\dagger}_{V,0}a^{\dagger}_{H,0}\\
&-\sqrt{\frac{1}{2}-\frac{5}{6\sqrt{3}}}a^{\dagger}_{H,0}a^{\dagger}_{V,0}a^{\dagger}_{H,-1}\Bigg)|0\rangle. \qquad (7)
\end{aligned}
$$

All the terms of the output state are divided into two parts: the contribution part and the non-contribution part. In the contribution part, the three photons occupy the three output ports $+1, 0, -1$, respectively. That is to say, there is only one photon in each of the three output ports $+1, 0, -1$, and the three photons are in the exact three-photon W state if we can postselect it from the output state. Other terms (abbreviated as $|\phi\rangle_{\text{other}}$) of the output state are categorized as the non-contribution part.

From the analysis of the output state, it is straightforward to see that three-photon W state can be generated by postselecting the event where there is only one photon in each of the three output ports $+1$, 0, -1 with probability $P = \left(\sqrt{\frac{2}{\sqrt{3}}-1}\right)^2 =$

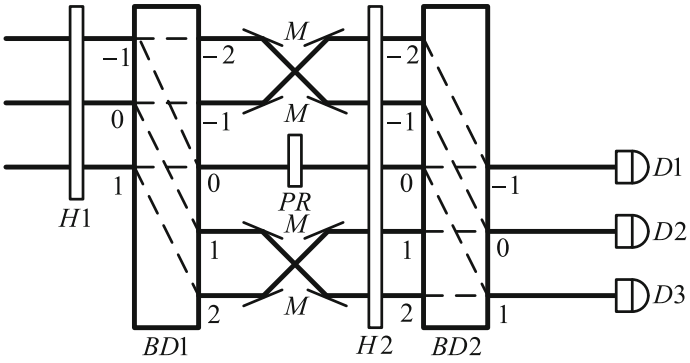


Fig. 2 The optical circuit for preparing polarization-entangled GHZ state

$\frac{2}{\sqrt{3}} - 1 \approx 15.47\%$. One can check that it is the maximum probability of obtaining W state, and it is higher than that of Ref. [49].

GHZ state is another important multipartite entangled state, which is inequivalent to W state, so the generation of GHZ state is definitely different from that of W state. To generate GHZ state via a similar mechanism used above for generating W state, we need to modify the optical circuit in Fig. 1 appropriately. The setup for generating GHZ state is depicted in Fig. 2, where two trajectory-exchanging operations are applied on the output modes $+1, +2$ and $-1, -2$ of $BD1$, respectively. In addition, the HWP $H1$ implements the operation $U1$ in Eq. (3). This setup will induce the system evolution from the input state in Eq. (1) to the following output state:

$$\begin{aligned}
|\psi^{(2)}\rangle &= \frac{1}{2\sqrt{2}}(a^\dagger_{H,1} + a^\dagger_{V,0})(a^\dagger_{H,-1} + a^\dagger_{V,1})(a^\dagger_{H,0} + a^\dagger_{V,-1})|0\rangle \\
&= \frac{1}{2}\left[\frac{1}{\sqrt{2}}(a^\dagger_{H,1}a^\dagger_{H,0}a^\dagger_{H,-1} + a^\dagger_{V,1}a^\dagger_{V,0}a^\dagger_{V,-1})|0\rangle\right] \\
&\quad + \frac{1}{2\sqrt{2}}(a^\dagger_{H,1}a^\dagger_{H,-1}a^\dagger_{V,-1} + a^\dagger_{H,1}a^\dagger_{V,1}a^\dagger_{H,0} + a^\dagger_{H,1}a^\dagger_{V,1}a^\dagger_{V,-1} \\
&\quad + a^\dagger_{V,0}a^\dagger_{H,-1}a^\dagger_{H,0} + a^\dagger_{V,0}a^\dagger_{H,-1}a^\dagger_{V,-1} + a^\dagger_{V,0}a^\dagger_{V,1}a^\dagger_{H,0})|0\rangle \\
&= \frac{1}{2}\left[\frac{1}{\sqrt{2}}(a^\dagger_{H,1}a^\dagger_{H,0}a^\dagger_{H,-1} + a^\dagger_{V,1}a^\dagger_{V,0}a^\dagger_{V,-1})|0\rangle\right] + |\phi\rangle_{\text{other}} \\
&= \frac{1}{2}|GHZ\rangle_3 + |\phi\rangle_{\text{other}},
\end{aligned} \tag{8}$$

where

$$\begin{aligned}
|\phi\rangle_{other} &= \frac{1}{2\sqrt{2}}(a^\dagger_{H,1}a^\dagger_{H,-1}a^\dagger_{V,-1} + a^\dagger_{H,1}a^\dagger_{V,1}a^\dagger_{H,0} + a^\dagger_{H,1}a^\dagger_{V,1}a^\dagger_{V,-1} \\
&\quad + a^\dagger_{V,0}a^\dagger_{H,-1}a^\dagger_{H,0} + a^\dagger_{V,0}a^\dagger_{H,-1}a^\dagger_{V,-1} + a^\dagger_{V,0}a^\dagger_{V,1}a^\dagger_{H,0})|0\rangle.
\end{aligned} \tag{9}$$

It is not difficult to see that three-photon GHZ state can be generated by postselecting the event where there is only one photon in each of the three output ports $+1$, 0, -1 with probability $P = (\frac{1}{2})^2 = \frac{1}{4}$.

After presenting detailed schemes for generating W and GHZ states, let's have a look at the physics underlying the schemes. There are two key points in our schemes. The first one is the indistinguishability of identical photons, which indicates that all the three photons are sharing the same spatial degree of freedom, and walking on the same beam displacer. That is so say, the three initially independent identical photons (by their spatial positions) get their spatial wave-functions overlapped during the protocol, so that at the output of the scheme one cannot determine which detected photon corresponds to which input photon. Mathematically, this means that the Hilbert spaces of each of the three photons are isomorphic and are spanned by the same set of vectors whose spatial degrees of freedom correspond to the values from -2 to 2. Formally, we are working in a three-particle sub-space of the overall photon-field Fock space, defined as a symmetric part of the tensor product of three (isomorphic) single-photon Hilbert spaces.

The second one is the trajectory-exchanging operation plus postselection. The output states in Eqs. (6) and (8) are rather complex entangled states, the component states can be divided into the contribution part and the non-contribution part, which is realized by the trajectory-exchanging operation. Then the contribution part can be easily singled out via the postselection. If there were no trajectory-exchanging operation, the output states couldn't be divided into the contribution part and the non-contribution part, and thus no entangled states could be singled out.

Up to now, we were studying the ideal case where after quantum walks the spatial wave-functions of the photons overlap perfectly on some positions of the beam displacers. What will happen if the spatial wave-functions of the photons do not overlap after trajectory exchanging (in general, this is what may happen in the lab)? If the spatial wave-functions of the photons do not overlap on some positions of the beam displacers, the photons become distinguishable (in time). Because all the derivations of this work are based on the identical condition of photons, if the photons become distinguishable, all the derivations of this paper do not hold anymore, and thus no entangled state will be generated in that case. Therefore, the temporal synchronicity problem will affect the schemes considerably, and the experimental setup must be designed in such a way that spatial wave-functions of different photons do overlap when they reach the same position of a beam displacer.

In addition, the advantage of using BDs rather than the standard PBSs is twofold. The first one is that the BD-based structure used here has the phase stability between all the different paths, but the PBS-based ones do not have this stability. This advantage has already been experimentally demonstrated [48]. The second one is that the BD-based structure can induce richer actions on the input photons than the PBS approach. Four BDs with trajectory exchanging can mimic, but not limited to, the action of a traditional PBS, because its function can be tuned if we insert some wave plates in between the four BDs. But the function of a PBS is fixed and cannot be tuned once it is manufactured.

Usually, the three input photons are generated from three different sources. To make these photons indistinguishable, we need the sophisticated on-demand single-photon

sources that have the features of superhigh efficiency and near-perfect purity and indistinguishabilty. Fortunately, efforts have been made to reach this target [66,67]. Therefore, we think our scheme will be feasible in the near future.

Finally, we would like to point out that, in order to generate the event-ready W and GHZ states using our scheme, an ideal "non-demolition" measurement of photons must be available. Although, within the current technology, it is impossible to detect a photon's presence without immediately annihilating it, we can combine the generation setup and its application setup together, and realize them simultaneously. Thus, at least we can avoid this problem in a feasible way.

3 Conclusion

In conclusion, optical schemes for generating photonic polarization-entangled W and GHZ states are proposed via quantum walk in this paper. The most distinct property of the schemes is that the coupling between photons is realized by trajectory-exchanging operation plus postselection, the simplicity of which greatly enhances its feasibility. Our schemes demonstrate that trajectory-exchanging operation plus postselection can single out entangled states from product states, which induces a non-dynamical coupling between two photons. The intrinsic physics underlying this new coupling mechanism is the indistinguishability of identical photons entering the device simultaneously.

In addition, the postselection operation and the trajectory-exchanging operation used here are not the only ones that can be realized. There are many different choices of them, which can induce more new multi-photon non-dynamical coupling mechanisms for quantum computation and quantum communication. Moreover, our new coupling mechanism for photons is more phase stable and tunable than the PBS-based one, which may greatly enrich the coupling mechanism for photons.

Acknowledgements This work is supported by the National Natural Science Foundation of China (NSFC)(11274010) and the personnel department of Anhui Province. NP acknowledges the support of SQIG–Security and Quantum Information Group, under the Project UID/EEA/50008/2019, IT Project QbigD funded by FCT PEst-OE/EEI/LA0008/2013 and the Project QuantumMining POCI-01-0145-FEDER-031826 funded by FCT through national funds, by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme (COMPETE 2020), and by Regional Operational Program of Lisbon.

References

1. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2000)
2. Walter, M., Grossy, D., Eisert, J.: Multi-partite entanglement. arXiv:1612.02437 (2016)
3. Dür, W., Vidal, G., Cirac, J.I.: Three qubits can be entangled in two inequivalent ways. Phys. Rev. A **62**(6), 062314 (2000)
4. Koashi, M., Bužek, V., Imoto, N.: Entangled webs: tight bound for symmetric sharing of entanglement. Phys. Rev. A **62**(5), R050302 (2000)
5. Dür, W.: Multipartite entanglement that is robust against disposal of particles. Phys. Rev. A **63**(2), R020303 (2001)

6. Bennett, C.H., Brassard, G.: Quantum cryptography: public key distribution and coin tossing. In: Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, New York, vol. 175, p. 8 (1984)
7. Ekert, A.K.: Quantum cryptography based on Bell's theorem. Phys. Rev. Lett. **67**(6), 661–663 (1991)
8. Bennett, C.H., Brassard, G., Crepeau, C., Jozsa, R., Peres, A., Wootters, W.K.: Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. Phys. Rev. Lett. **70**(13), 1895–1899 (1993)
9. Bouwmeester, D., Pan, J.W., Mattle, K., Eibl, M., Weinfurter, H., Zeilinger, A.: Experimental quantum teleportation. Nature **390**, 575–579 (1997)
10. Yeo, Y., Chua, W.K.: Teleportation and dense coding with genuine multipartite entanglement. Phys. Rev. Lett. **96**(6), 060502 (2006)
11. Bennett, C.H., Wiesner, S.J.: Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states. Phys. Rev. Lett. **69**(20), 2881–2884 (1992)
12. Zhao, Z., Chen, Y.A., Zhang, A.N., Yang, T., Briegel, H.J., Pan, J.W.: Experimental demonstration of five-photon entanglement and open-destination teleportation. Nature **430**, 54–58 (2004)
13. Kempe, J.: Multiparticle entanglement and its applications to cryptography. Phys. Rev. A **60**(2), 910–916 (1999)
14. Ma, J., Huang, Y.X., Wang, X.G., Sun, C.P.: Quantum fisher information of the Greenberger–Horne–Zeilinger state in decoherence channels. Phys. Rev. A **84**(2), 022302 (2011)
15. Greenberger, D.M., Horne, M.A., Zeilinger, A.: Going beyond Bell's theorem. In: Kafatos, M. (ed.) Bell's Theorem, Quantum Theory, and Conceptions of the Universe, p. 69. Kluwer, Dordrecht (1989)
16. Mermin, N.D.: Extreme quantum entanglement in a superposition of macroscopically distinct states. Phys. Rev. Lett. **65**(15), 1838–1840 (1990)
17. Pan, J.W., Bouwmeester, D., Daniell, M., Weinfurter, H., Zeilinger, A.: Experimental test of quantum nonlocality in three-photon Greenberger–Horne–Zeilinger entanglement. Nature **403**, 515–519 (2000)
18. Häffner, H., Hänsel, W., Roos, C.F., Benhelm, J., Chek-al-kar, D., Chwalla, M., Körber, T., Rapol, U.D., Riebe, M., Schmidt, P.O., Becher, C., Gühne, O., Dür, W., Blatt, R.: Scalable multiparticle entanglement of trapped ions. Nature **438**, 643–646 (2005)
19. Monz, T., Schindler, P., Barreiro, J.T., Chwalla, M., Nigg, D., Coish, W.A., Harlander, M., Hänsel, W., Hennrich, M., Blatt, R.: 14-Qubit entanglement: creation and coherence. Phys. Rev. Lett. **106**(13), 130506 (2011)
20. Song, C., Xu, K., Liu, W., Yang, C.P., Zheng, S.B., Deng, H., Xie, Q., Huang, K., Guo, Q., Zhang, L., Zhang, P., Xu, D., Zheng, D., Zhu, X., Wang, H., Chen, Y.A., Lu, C.Y., Han, S., Pan, J.W.: 10-Qubit entanglement and parallel logic operations with a superconducting circuit. Phys. Rev. Lett. **119**(18), 180511 (2017)
21. Neeley, M., Bialczak, R.C., Lenander, M., Lucero, E., Mariantoni, M., O'Connell, A.D., Sank, D., Wang, H., Weides, M., Wenner, J., Yin, Y., Yamamoto, T., Cleland, A.N., Martinis, J.M.: Generation of three-qubit entangled states using superconducting phase qubits. Nature **467**, 570–573 (2010)
22. DiCarlo, L., Reed, M.D., Sun, L., Johnson, B.R., Chow, J.M., Gambetta, J.M., Frunzio, L., Girvin, S.M., Devoret, M.H., Schoelkopf, R.J.: Preparation and measurement of three-qubit entanglement in a superconducting circuit. Nature **467**, 574–578 (2010)
23. Guo, G.C., Zhang, Y.S.: Scheme for preparation of the W state via cavity quantum electrodynamics. Phys. Rev. A **65**(5), 054302 (2002)
24. Guo, G.P., Li, C.F., Li, J., Guo, G.C.: Scheme for the preparation of multiparticle entanglement in cavity QED. Phys. Rev. A **65**(4), 042102 (2002)
25. Xue, P., Guo, G.C.: Scheme for preparation of mulipartite entanglement of atomic ensembles. Phys. Rev. A **67**(3), 034302 (2003)
26. Xue, P., Han, C., Yu, B., Lin, X.M., Guo, G.C.: Entanglement preparation and quantum communication with atoms in optical cavities. Phys. Rev. A **69**(5), 052318 (2004)
27. Gao, Y.X., Zhou, H., Zou, D., Peng, X.H., Du, J.F.: Preparation of Greenberger–Horne–Zeilinger and W states on a one-dimensional Ising chain by global control. Phys. Rev. A **87**(3), 032335 (2013)
28. Chen, J.H., Zhou, H., Duan, C.K., Peng, X.H.: Preparing Greenberger–Horne–Zeilinger and W states on a long-range Ising spin model by global controls. Phys. Rev. A **95**(3), 032340 (2017)
29. Yamamoto, T., Tamaki, K., Koashi, M., Imoto, N.: Polarization-entangled W state using parametric down-conversion. Phys. Rev. A **66**(6), 064301 (2002)
30. Greganti, C., Roehsner, M.C., Barz, S., Waegell, M., Walther, P.: Practical and efficient experimental characterization of multiqubit stabilizer states. Phys. Rev. A **91**(2), 022325 (2015)

31. Ikuta, R., Tashima, T., Yamamoto, T., Koashi, M., Imoto, N.: Optimal local expansion of W states using linear optics and Fock states. Phys. Rev. A **83**(1), 012314 (2011)
32. Fujii, K., Maeda, H., Yamamoto, K.: Robust and scalable scheme to generate large-scale entanglement webs. Phys. Rev. A **83**(5), R050303 (2011)
33. Özdemir, Ş.K., Matsunaga, E., Tashima, T., Yamamoto, T., Koashi, M., Imoto, N.: An optical fusion gate for W-states. New J. Phys. **13**, 103003 (2011)
34. Lu, C.Y., Zhou, X.Q., Gühne, O., Gao, W.B., Zhang, J., Yuan, Z.S., Goebel, A., Yang, T., Pan, J.W.: Experimental entanglement of six photons in graph states. Nat. Phys. **3**, 91–95 (2007)
35. Bugu, S., Yesilyurt, C., Ozaydin, F.: Enhancing the W-state quantum-network-fusion process with a single Fredkin gate. Phys. Rev. A **87**(3), 032331 (2013)
36. Ozaydin, F., Bugu, S., Yesilyurt, C., Altintas, A.A., Tame, M., Özdemir, Ş.K.: Fusing multiple W states simultaneously with a Fredkin gate. Phys. Rev. A **89**(4), 042311 (2014)
37. Zang, X.P., Yang, M., Ozaydin, F., Song, W., Cao, Z.L.: Generating multi-atom entangled W states via light-matter interface based fusion mechanism. Sci. Rep. **5**, 16245 (2015)
38. Zang, X.P., Yang, M., Song, W., Cao, Z.L.: Fusion of entangled coherent W and GHZ states in cavity QED. Opt. Commun. **370**, 168–171 (2016)
39. Zang, X.P., Yang, M., Ozaydin, F., Song, W., Cao, Z.L.: Deterministic generation of large scale atomic Wstates. Opt. Express **24**(11), 12293–12230 (2016)
40. Li, K., Kong, F.Z., Yang, M., Yang, Q., Cao, Z.L.: Qubit-loss-free fusion of W states. Phys. Rev. A **94**(6), 062315 (2016)
41. Wang, X.L., Chen, L.K., Li, W., Huang, H.L., Liu, C., Chen, C., Luo, Y.H., Su, Z.E., Wu, D., Li, Z.D., Lu, H., Hu, Y., Jiang, X., Peng, C.Z., Li, L., Liu, N.L., Chen, Y.A., Lu, C.Y., Pan, J.W.: Experimental ten-photon entanglement. Phys. Rev. Lett. **117**(21), 210502 (2016)
42. Wang, X.L., Luo, Y.H., Huang, H.L., Chen, M.C., Su, Z.E., Liu, C., Chen, C., Li, W., Fang, Y.Q., Jiang, X., Zhang, J., Li, L., Liu, N.L., Lu, C.Y., Pan, J.W.: 18-Qubit entanglement with six photons' three degrees of freedom. Phys. Rev. Lett. **120**(26), 260502 (2018)
43. Paunković, N.: The role of indistiguishability of identical particles in quantum information processing, Ph.D. thesis, Keble College, University of Oxford (2004). URL = http://sqig.math.ist.utl.pt/pub/PaunkovicN/04-P-phdthesis.pdf
44. Lo Franco, R., Compagno, G.: Quantum entanglement of identical particles by standard information-theoretic notions. Sci. Rep. **6**, 20603 (2016)
45. Lo Franco, R.: Compagno: indistinguishability of elementary systems as a resource for quantum information processing. Phys. Rev. Lett. **120**(24), 240403 (2018)
46. Yurke, B., Stoler, D.: Einstein–Podolsky–Rosen effects from independent particle sources. Phys. Rev. Lett. **68**(9), 1251–1254 (1992)
47. Yurke, B., Stoler, D.: Bell's-inequality experiments using independent-particle sources. Phys. Rev. A **46**(5), 2229–2234 (1992)
48. Kim, Y.-S., Pramanik, T., Cho, Y.-W., Yang, M., Han, S.-W., Lee, S.-Y., Kang, M.-S., Moon, S.: Informationally symmetrical Bell state preparation and measurement. Opt. Express **26**(22), 29539 (2018)
49. Blasiak, P., Markiewicz, M.: Entangling three qubits without ever touching. arXiv:1807.05546 (2018)
50. Li, X.M., Yang, M., Paunković, N., Li, D.C., Cao, Z.L.: Entanglement swapping via three-step quantum walk-like protocol. Phys. Lett. A **381**(46), 3875–3879 (2017)
51. Aharonov, Y., Davidovich, L., Zagury, N.: Quantum random walks. Phys. Rev. A **48**(2), 1687–1690 (1993)
52. Kempe, J.: Quantum random walks—an introductory overview. Contemp. Phys. **44**(4), 307–327 (2003)
53. Xue, P., Sanders, B.C., Leibfried, D.: Quantum walk on a line for a trapped ion. Phys. Rev. Lett. **103**(18), 183602 (2009)
54. Xue, P., Sanders, B.C.: Two quantum walkers sharing coins. Phys. Rev. A **85**(2), 022307 (2012)
55. Xue, P., Zhang, R., Bian, Z.H., Zhan, X., Qin, H., Sanders, B.C.: Localized state in a two-dimensional quantum walk on a disordered lattice. Phys. Rev. A **92**(4), 042316 (2015)
56. Karski, M., Förster, L., Choi, J.M., Steffen, A., Alt, W., Meschede, D., Widera, A.: Quantum walk in position space with single optically trapped atoms. Science **325**(5937), 174–177 (2009)
57. Broome, M.A., Fedrizzi, A., Lanyon, B.P., Kassal, I., Aspuru-Guzik, A., White, A.G.: Discrete single-photon quantum walks with tunable decoherence. Phys. Rev. Lett. **104**(15), 153602 (2010)
58. Kurzyński, P., Wójcik, A.: Quantum walk as a generalized measuring device. Phys. Rev. Lett. **110**(20), 200404 (2013)

59. Bian, Z.H., Li, J., Qin, H., Zhan, X., Zhang, R., Sanders, B.C., Xue, P.: Realization of single-qubit positive-operator-valued measurement via a one-dimensional photonic quantum walk. Phys. Rev. Lett. **114**(20), 203602 (2015)
60. Hou, Z.B., Tang, J.F., Shang, J.W., Zhu, H.J., Li, J., Yuan, Y., Wu, K.D., Xiang, G.Y., Li, C.F., Guo, G.C.: Deterministic realization of collective measurements via photonic quantum walks. Nat. Commun. **9**, 1414 (2018)
61. Shenvi, N., Kempe, J., Whaley, K.B.: Quantum random-walk search algorithm. Phys. Rev. A **67**(5), 052307 (2003)
62. Childs, A.M.: Universal computation by quantum walk. Phys. Rev. Lett. **102**(18), 180501 (2009)
63. Lovett, N.B., Cooper, S., Everitt, M., Trevers, M., Kendon, V.: Universal quantum computation using the discrete-time quantum walk. Phys. Rev. A **81**(4), 042330 (2010)
64. Childs, A.M., Gosset, D., Webb, Z.: Universal computation by multiparticle quantum walk. Science **339**(6121), 791–794 (2013)
65. Wang, Y., Shang, Y., Xue, P.: Generalized teleportation by quantum walks. Quant. Inf. Proc. **16**(9), 221 (2017)
66. Ding, X., He, Y., Duan, Z.C., Gregersen, N., Chen, M.C., Unsleber, S., Maier, S., Schneider, C., Kamp, M., Höfling, S., Lu, C.Y., Pan, J.W.: On-demand single photons with high extraction efficiency and near-unity indistinguishability from a resonantly driven quantum dot in a micropillar. Phys. Rev. Lett. **116**(2), 020401 (2016)
67. He, Y.M., Liu, J., Maier, S., Emmerling, M., Gerhardt, S., Davanço, M., Srinivasan, K., Schneider, C., Höfling, S.: Deterministic implementation of a bright, on-demand single-photon source with near-unity indistinguishability via quantum dot imaging. Optica **4**(7), 802–808 (2017)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Quantum Information Processing (2018) 17:288
https://doi.org/10.1007/s11128-018-2055-y

Quantum key distribution with quantum walks

Chrysoula Vlachou[1,2] · Walter Krawec[4] · Paulo Mateus[1,2] · Nikola Paunković[1,2,3] · André Souto[1,5,6]

Received: 18 December 2017 / Accepted: 6 September 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract
Quantum key distribution is one of the most fundamental cryptographic protocols. Quantum walks are important primitives for computing. In this paper, we take advantage of the properties of quantum walks to design new secure quantum key distribution schemes. In particular, we introduce a secure quantum key distribution protocol equipped with verification procedures against full man-in-the-middle attacks. Furthermore, we present a one-way protocol and prove its security. Finally, we propose a semi-quantum variation and prove its robustness against eavesdropping.

Keywords Quantum walks · Quantum key distribution · Semi-quantum key distribution

1 Introduction

Quantum Key Distribution (QKD) is the most secure and practical instance of quantum cryptography. We recall that a key distribution scheme is a protocol between two parties with the purpose of sharing a common string (the key), which afterward, they can use to communicate privately, in a pre-agreed encryption scheme. Therefore, it is required that any third party that might be eavesdropping is not able to extract information about the key, and thus compromising the privacy of the communication. Bennet and Brassard [10] in 1984, and Ekert [33] in 1991, proposed the first QKD protocols, upon which all QKD protocols are based. Since then a lot of modifications and improvements have been proposed in order to achieve unconditionally secure and practical QKD schemes, by taking advantage of the physical laws of quantum mechanics. For a review, see [96]. Several QKD experiments over long distances have

✉ André Souto
asouto@fc.ul.pt

Extended author information available on the last page of the article

Published online: 15 September 2018

been reported [83,85,102,103], and QKD is already commercial.[1] Furthermore, the recent successful launch of a satellite [116] paved the way for intercontinental QKD.

Quantum walks (QW) have been introduced in 1993, in [2], as the quantum analogue of classical random walks. Since then, they have been playing a major role in quantum computing theory, as their applications vary from quantum algorithms [3,23,36,82,94] to universal quantum computing schemes [24,25,65].

Recently, the application of QWs to the creation of actual quantum cryptographic protocols has been investigated. For instance, in [89], Rohde et al. proposed a limited form of quantum homomorphic encryption using multi-particle QWs. In their protocol, a server could manipulate data sent by a client in such a way that, first, the server has limited information on the client's data while, second, the client has limited information on the server's computation.

In this paper, we revisit the public-key cryptosystem [106], which is based on QWs, in order to construct secure QKD protocols. First, we suitably modify [106], so that the quantum state generated by means of a QW encodes the secret key as opposed to the message; such a key could be used later as input to a one-time-pad encryption system gaining information theoretic security for message delivery. Our motivation is that QKD schemes have several advantages, which we present in due course, over public-key cryptosystems. The modification of the original public-key system is non-trivial, however, and requires care as we can no longer rely on the existence of a trusted mechanism for public-key delivery (such as a public-key infrastructure), as is typically assumed in quantum public-key cryptography [79,99,106].

While the above QKD protocol is two-way, i.e., both Alice (A) and Bob (B) perform QW operations, we also construct a one-way QKD protocol, where again the key is encoded in a QW state. In this case, it is only Alice that chooses randomly the precise QW to encode the key, while Bob is randomly choosing in which basis (computational or QW) to measure in order to obtain it. After disclosing their choices by means of classical communication, they are able to establish a shared key. We prove that the protocol is secure against general attacks, even if the eavesdropper Eve (E) has great advantage over Alice and Bob.

As a third contribution in this paper, we propose a new *semi-quantum* key distribution (SQKD) protocol based on QWs. Semi-quantum cryptography was first introduced in 2007 by Boyer et al. [13,14] as a way to study "how quantum" does a protocol need to be in order to gain an advantage over its classical counterpart—namely, how quantum do the parties need to be in order to establish a secret key secure against an all powerful adversary. Using classical communication alone, this task is impossible—indeed, any key distribution protocol, relying only on classical communication, cannot be unconditionally secure and, instead, requires computational hardness assumptions to be made on the adversary. On the other hand, QKD protocols do have provable unconditional security. A semi-quantum protocol places severe restrictions on one of the participating users (typically Bob) in that he may only operate in a "classical" or "semi-quantum" manner. Namely, this limited user can only directly work with the

[1] Currently, there are three companies offering commercial QKD systems: ID Quantique (Geneva), MagiQ Technologies, Inc. (New York) and QuintessenceLabs (Australia).

computational Z basis. No restrictions are placed on the other participant Alice, and of course, no restrictions are placed on Eve.

The paper is organized as follows: In Sect. 2, we provide a brief introduction to QWs, with all the information and notation used throughout the paper. In Sect. 3, we present a secure two-way QKD scheme based on QWs, which is a modification of the public-key cryptosystem in Ref. [106]. We give motivation for this modification, and furthermore, we propose two different verification procedures against full man-in-the-middle attacks. In Sect. 4, we introduce a one-way QKD protocol, which we prove to be unconditionally secure, by reducing it to an equivalent entanglement-based protocol. We provide our numerical results for the optimal choice of the QW parameters that maximize the noise tolerance of the protocol. Finally, in Sect. 5, we provide a SQKD protocol and we show its robustness against eavesdropping.

2 Quantum walk preliminaries

In this paper, we consider QWs on a circle. In this case, the walker hops along discrete positions on a circle. The Hilbert space $\mathcal{H}$, describing the QW, is the tensor product of the positions Hilbert space $\mathcal{H}_p$ and the coin Hilbert space $\mathcal{H}_c$, i.e., $\mathcal{H} = \mathcal{H}_p \otimes \mathcal{H}_c$. The positions Hilbert space is spanned by the points on a circle $\{|x\rangle \, | x \in \{0, \ldots, P-1\}\}$, while $\mathcal{H}_c$ is spanned by the two possible coin states $\{|R\rangle, |L\rangle\}$, corresponding to heads and tails. The evolution for one step of the QW is given by the unitary operator

$$U = S \cdot (I_p \otimes R_c)$$

where I_p is the identity operator in $\mathcal{H}_p$, $R_c \in SU(2)$ is a rotation in $\mathcal{H}_c$, which in matrix form we can write it as:

$$R_c(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}, \tag{1}$$

and

$$S = \sum_{x=0}^{P-1} \Big[|x+1 \pmod P\rangle \langle x| \otimes |R\rangle \langle R| \tag{2}$$

$$+ |x-1 \pmod P\rangle \langle x| \otimes |L\rangle \langle L| \Big] \tag{3}$$

is the shift operator that moves the walker one position to the right or to the left on the circle, depending on its coin state. Notice that, since we are on a circle, the P-th position is identified with the 0 position.

3 Quantum walk key distribution scheme

In this section, we introduce a QKD scheme based on QWs. In this context, the key for the QKD is encrypted as the message in the public-key cryptosystem introduced in [106]. This modification is motivated by the fact that QKD schemes are more flexible than public-key protocols, as the key can be used by both Alice and Bob to send or authenticate messages. Also, more post-processing techniques (e. g. privacy amplification) can be applied, since we have as input a random string and not a plaintext message. In the latter case, we should be careful during the post-processing not to degrade the message (we are left with less techniques). Furthermore, in the case of information leakage, we can safely abort the protocol, while during message transmission it would be late for that.

Our QW-QKD scheme is depicted in Fig. 1 and presented below. We assume that the key can be chosen among P possible keys. We also assume that the QW can be chosen from a prefixed discrete set known by both parties.

Protocol 1 *Quantum key distribution scheme*

Inputs for the protocol

- *Key:*
  $r \in \{0, \ldots, P-1\}$*, i.e., a key of at most* $\log P$ *bits, chosen by Bob uniformly at random;*
- *Quantum state generation:*
  The QW operator U_k *with* $k \in \mathcal{K} = \{1, 2, \ldots, K\}$*, the number of steps* $t \in \mathcal{T} = \{T_0, \ldots, T_{max}\} \subset \mathbb{N}$*, and the initial state* $|l\rangle \otimes |s\rangle$*, where* $l \in \{0, \ldots, P-1\}$, $s \in \{R, L\}$.

In the above, U_k*, the QW operator is defined as* $U_k = S \cdot (I_p \otimes R_c(\theta_k))$*, where* S *is the shift operator and* $R_c(\theta_k)$ *is a rotation of* $\theta_k = k \cdot 2\pi/K$ *in the coin space, see Eqs. 1 and 2.*

Quantum state generation

- *Alice chooses uniformly at random* $l \in \{0, \ldots, P-1\}$ *and* $s \in \{R, L\}$*, and generates the initial state* $|l\rangle\,|s\rangle$.

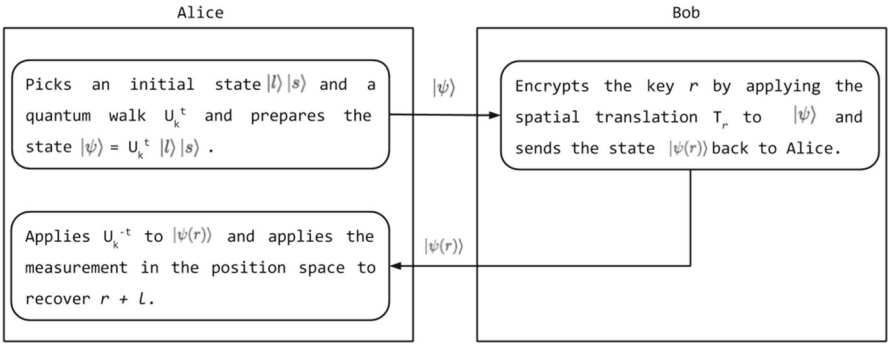


Fig. 1 Description of the basic steps of Protocol 1

– *Then, she chooses, also at random, the QW* $U_k = S \cdot \left(I_p \otimes R_c(\theta_k)\right)$ *and the number of steps* $t \in \mathcal{T}$.
– *Finally, she generates the quantum state:*

$$|\psi\rangle = U_k^t \, |l\rangle \, |s\rangle = \left[S \cdot \left(I_p \otimes R_c(\theta_k)\right)\right]^t |l\rangle \, |s\rangle \, ,$$

and sends it to Bob.

Key encryption

– *Upon obtaining the quantum state* $|\psi\rangle$ *from Alice, Bob encrypts the key* r *by applying spatial translation* $T_r = \sum_{i=0}^{P-1} |i + r \pmod P\rangle \langle i|$ *to obtain:*

$$|\psi(r)\rangle = (T_r \otimes I_c) \, |\psi\rangle \, ,$$

where I_c *is the identity operator in the coin space.*
– *Bob sends* $|\psi(r)\rangle$ *to Alice.*

Key decryption

– *Alice applies* U_k^{-t} *to the state* $|\psi(r)\rangle$.
– *She performs the position measurement*

$$M = \sum_{i=0}^{P-1} |i\rangle \langle i| \otimes I_c$$

and obtains the result i_0.
The key sent by Bob is $r = i_0 - l \pmod P$.

It is clear, from the design of the protocol and the proof of correctness of the original quantum encryption scheme presented in [106] that, if no one interferes with the quantum states, then the protocol is correct and at the end, Alice and Bob will share a common string of length $\log P$, that they can use as a key. In the next section, we prove the security of the protocol.

3.1 Security of the Scheme

In [106], the authors use the Holevo theorem to show that Eve can extract information about the key, by means of the quantum states $|\psi\rangle$ and $|\psi(r)\rangle$ that Alice and Bob exchange, only with negligible probability. Here, we do not present this proof of security for the sake of briefness, but the interested reader can find it in [106] with all the details. There, it was shown that the protocol is secure if the size of the space of parameters, which are chosen uniformly at random, is exponentially large with respect to the length of the message transmitted (the key in our case of a QKD protocol), see Equation (19) and the comment below it, in Sect. 3.2 of [106]. Moreover, for the protocol to be efficient, it was shown that the size of the $\mathcal{T}$ (set of possible number of steps) should be polynomial with respect to the length of the message/key transmitted,

see Section 3.3 of [106]. Therefore, to maintain the protocol's security, it is necessary that the size of $\mathcal{K}$, the set of possible coin unitaries U_k, is exponentially large with respect to the size of the key. As a consequence, while it is possible to fix the number of steps of the QW, and thus somewhat simplify its implementation, to maintain the security, it is necessary to keep the parameter k of the protocol.

Another type of attack that Eve can perform is a full man-in-the-middle attack, in which she impersonates Alice to Bob and vice versa, while they think that they are communicating directly. This attack gives Eve the chance to intercept and alter the communication between them. In public-key cryptosystems, such attacks can be prevented by using a public-key infrastructure, which is assumed to work as a trusted third party. In our QKD modification though, such an assumption could not be used; therefore, we should complete the security analysis of the scheme, by taking into account full man-in-the-middle attacks. To this end, we propose two different verification procedures, that allow Alice and Bob to verify that what they receive is actually coming from each other and not from an eavesdropper pretending to be either of them. We should note that for both verification methods, Alice and Bob need to share a classical public authenticated channel (a common requirement in QKD protocols, such as the well-known BB84 scheme [10]).

3.1.1 Standard verification

The first technique we propose is a standard cut-and-choose verification, which is achieved by adding redundancy to our scheme. Clearly, the verification is needed twice in our protocol: once when Alice sends the QW state to Bob and once when Bob sends the encoded key to Alice.

Verification 1: *Bob verifies that it was Alice who sent him the quantum state.*

It is needed to prevent Eve from sending her choice of quantum states to Bob, which would allow her to read the encrypted key while he is sending it back to Alice.

- Alice sends to Bob $\bigotimes_{i=1}^{m} |\psi_i\rangle$, that is, several quantum states $|\psi_i\rangle$, generated by a QW as described in the previous section. Each $|\psi_i\rangle$ is generated using independently chosen walk parameters and initial states (k_i, t_i, l_i, s_i).
- After Bob receiving $\bigotimes_{i=1}^{m} |\psi_i\rangle$, Alice, through a classical authenticated channel, sends him a string $v = v_1 v_2 \dots v_m$ of m bits, such that $v_i = 1$ if the corresponding $|\psi_i\rangle$ is going to be used for verification and $v_i = 0$ otherwise, that is, if the corresponding $|\psi_i\rangle$ will be used by Bob to encode part of the key. Through the classical channel, she also sends (j, k_j, t_j, l_j, s_j), for some uniformly at random chosen j's that belong in the set $\{1, \dots, m\}$. Let the number of these j's be $m/3$.
- Bob verifies that for all these j's, the received states $\rho_j = |\psi_j\rangle \langle\psi_j|$ are indeed equal to the pure states

$$|\psi_j\rangle = U_{k_j}^{t_j} |l_j\rangle |s_j\rangle .$$

  In order to verify that, he applies $U_{k_j}^{-t_j}$ to the states $|\psi_j\rangle$, for all j and then performs a measurement for each j in the positions space as well as in the spin space. This measurement (for each j) is described by the operator:

$$M_{l_j,s_j} = \sum_{l_j,s_j} \alpha_{l_j,s_j} |l_j, s_j\rangle \langle l_j, s_j| = \sum_{l_j} l_j |l_j\rangle \langle l_j| \otimes \sum_{s_j} s_j |s_j\rangle \langle s_j| .$$

This way, he traces out all these $|\psi_j\rangle$'s and he is left with $2m/3$ quantum states. We call the reader's attention to the fact that if the verification fails for any j, the protocol is stopped.

Verification 2: *Alice verifies that it was Bob who sent her the encrypted key.*

This procedure is needed to prevent Eve from sending to Alice a message that would decrypt a key different from the one sent by Bob. In this case, Alice and Bob would not be able to communicate, while Eve would be able to decrypt messages sent by Alice (not vice versa). To prevent this from happening, Alice and Bob repeat the verification procedure 1, with the roles switched. In particular, the two are performing the following steps:

- Bob encrypts $r_i \in \{0, \ldots, P-1\}$ in each of the states of the remaining product state $\bigotimes_{i=1}^{2m/3} |\psi_i\rangle$ as follows:

$$|\psi(r_i)\rangle = (T_{r_i} \otimes I_c) |\psi_i\rangle , \forall i \in \{1, \ldots, 2m/3\}$$

  that is, translating each state $|\psi_i\rangle$ by r_i in the positions space, leaving the spin part of the state unaltered.
- He sends the product state $\bigotimes_{i=1}^{2m/3} |\psi(r_i)\rangle$ to Alice.
- Then, he chooses $m/3$ uniformly at random j''s out of the $2m/3$ unused indices from the previous verification procedure. Through the classical public authenticated channel, he sends a classical string $v' = v'_1 v'_2 \ldots v'_{2m/3}$ of $2m/3$ bits, such that $v'_i = 1$ if the corresponding $|\psi(r_i)\rangle$ is going to be used for verification and $v'_i = 0$ otherwise, that is, if the corresponding $|\psi(r_i)\rangle$ contains part of the key. For each j' chosen (for which $v'_i = 1$), he also sends through the classical public authenticated channel the index and the respective $r_{j'}$'s used to generate the state $|\psi(r_i)\rangle$.
- In the last step, Alice applies $U_{k_i}^{-t_i}$ on the $2m/3$ states $|\psi(r_i)\rangle$ and then, for each i, she performs a measurement on the positions space. Let the outcomes be denoted by $\alpha_i, i \in \{1, \ldots, 2m/3\}$. For all the indices, she computes $r_i = \alpha_i - l_i$, where l_i are the initial positions on the circle that she used for the generation of the quantum states $|\psi_i\rangle$. Finally, for each $j', r_{j'}$ sent by Bob, Alice verifies the consistency of their results.
- The key is given by the concatenation of the bits r_i that were not used during the two verification procedures and it has $m \cdot (\log P)/3$ bits. Usually, the choice of m is dependent on the desired length, $\log P$, of the key, and in order to make the success probability of a man-in-the-middle attack negligible on $\log P$, it is common to use $m = (\log P)/3$.

3.1.2 Verification using maximally entangled states

In this section, we present an alternative verification procedure, which prevents Eve from trying to infer the key by first entangling her ancillas with the systems sent by Alice, and then performing an additional operation (say, a measurement) on the joint system of her ancillas and those carrying the encrypted key sent back to Alice by Bob; a method which in general would give her access to some non-negligible amount of information, so that Alice and Bob are not able to securely communicate. Note that this verification procedure could also be used against the previous attack in which Eve simply impersonates Alice to Bob, and vice versa.

During the first step of the protocol ("Quantum state generation"), in addition to generating QW states

$$|\psi\rangle_{qw} = U_k^t |l\rangle |s\rangle \tag{4}$$

used to encode the key, for the verification purposes, Alice also creates a number of Bell-like maximally entangled states

$$|\psi\rangle_{qw} = \frac{1}{\sqrt{(\log 2P)!}} \sum_{i=0}^{2P-1} |i\rangle_a |i\rangle_{qw} \tag{5}$$

between the ancilla systems (denoted by a) and the QW systems (denoted by qw), each of dimension $2P$ (the dimension of the actual QW). At the end of the first step, Alice sends to Bob a random sequence of QW states, each either in the form $|\psi\rangle_{qw}$, or $\rho_{qw} = \mathrm{Tr}_a |\psi\rangle \langle\psi|_{qw}$, while keeping the ancillas with her. Alice also sends through a classical public authenticated channel a classical string $v = v_1 \ldots v_n$, where $v_i = 0$ if the i-th system is going to be used for the encoding of the key, while $v_i = 1$ if the i-th system is going to be used for verification.

The proportion of states used to obtain the key and used for the verification can be chosen in a similar way as in the previous case. Usually, the dimension of the total Hilbert space $2P$ is of the form 2^n which, in turn, is isomorphic to the Hilbert space resulting from the tensor product of n 2-dimensional Hilbert spaces, and thus, this state can be written as the tensor product of n standard two-qubit $|\phi^+\rangle$ Bell states.

After Bob receives the systems, he and Alice perform Bell-like measurements on the states meant for the verification and they observe a maximal violation of the Bell's inequalities, since those states are maximally entangled. This way, these states are traced out and Bob is left with the states (4) in which he will encode the key (as previously).

The same procedure is repeated again, when Bob sends the encoded key to Alice. He will send a sequence of states, some of the form $(\hat{T}_r \otimes \hat{I}_c) U_k^t |l\rangle |s\rangle$, in which part of the key is encoded and some of the form (4) (with his ancillary system $|i\rangle_B$ maximally entangled to the system sent to Alice), which are going to be used for the verification, as explained above. In the end of the key decryption phase and if all the verifications were okay, Alice will concatenate the parts of the key to obtain the full key.

3.2 Efficiency and quantum memory requirements

In [106], it has been proven that the protocol is efficient, i.e., it requires only polynomial time (on the length of the message, say n) to transfer n bits of information encoded in $n+1$ qubits. By introducing the verification steps in this QKD scheme, we increase the complexity of the system to n^2, in order to make the probability of eavesdropping negligible. However, we should notice that, out of this scheme, the size of the key that Alice and Bob share at the end is also increased to $n^2/3$, considering $m = n$. Therefore, the number of bits in the key is linear in the number of qubits sent to Bob. As a conclusion, our QKD scheme is efficient, since the complexity increased, but only polynomially.

As already mentioned in the Introduction, the lack of stable quantum memories is a major issue in quantum cryptography, since it is a practical constraint that is not likely to be solved, at least in the near future. Short-term quantum memories already exist; however, it is not always straightforward to argue about the security of a protocol, relying on their existence. In our case, though, things are quite clear. If Eve does not interfere, Alice and Bob do not need quantum memories to execute the protocol; thus, the key distribution is independent of such practical constraints. However, the presence of Eve and the need of verification for Alice and Bob introduce memory requirements for *all* the parties.

Below, we present the memory requirements for the case of Sect. 3.1.1, noting that the case of Sect. 3.1.2 is analogous. To conduct her attack, Eve needs a stable quantum memory, in order to keep the states she intercepted by Alice, while waiting for Bob to encrypt and send the key. Subsequently, she will encode it in Alice's states and send it to her. Also, in this scenario, Alice and Bob need a quantum memory, in order to perform the verification. They need to save the quantum states for some time, while waiting for the other party to send the classical information. Observe that Eve's memory should be more stable than Alice's and Bob's, as the time Eve needs to save the quantum states for, is clearly longer than the time that Alice and Bob need for the same purpose.

Hence, we conclude that our QKD scheme is secure, as long as Alice and Bob have at least as powerful equipment as the adversary Eve. Obviously, if the adversary is technologically more advanced, then virtually any real-life implementation of a security protocol becomes potentially vulnerable.

4 One-way quantum walk key distribution protocol

In this section, we present a one-way QKD protocol based on QWs and we prove that it is secure. First, we state it in its prepare-and-measure form. The protocol procedure is depicted in Fig. 2.

Protocol 2 *Let θ_k, t, and P be publicly known where P is the dimension of the position space of the QW, t is the number of steps to perform the QW, and θ_k the coin parameter [see Eq. ((1))]. Let U_k be the QW operator $U_k = S \cdot (I_p \otimes R_c(\theta_k))$ that is also known by the parties (i.e., it is also publicly known) and let F be an operator acting only on*

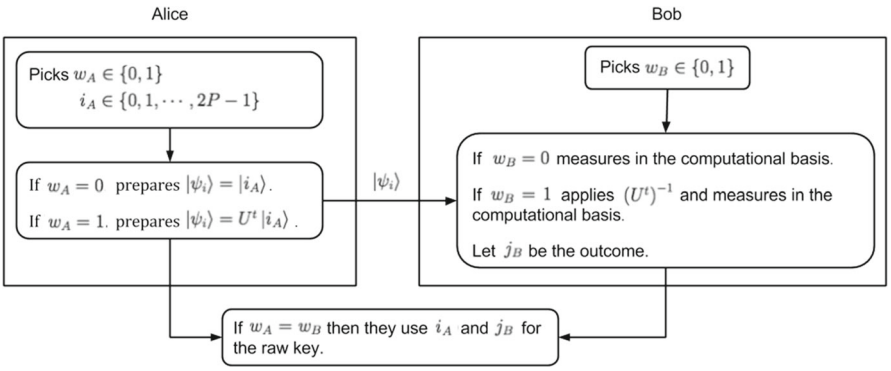


Fig. 2 Description of the basic steps of Protocol 2

$\mathcal{H}_c$. *F's action is to "flip" the coin to some initial state before evolving the walk and is optional (in which case* $F = I_c$*). Finally, let* $|\psi_i\rangle = U_k^t(I_p \otimes F)\,|i\rangle$ *for* $|i\rangle \in \mathcal{H}_p \otimes \mathcal{H}_c$. *We call the orthonormal basis* $\{|\psi_i\rangle\}$ *the QW basis and denote the computational basis by Z. This means that the QW basis is obtained from the computational basis Z, when performing the QW with respect to the initial state* $|i\rangle$ *or, in other words, the unitary operator that describes the basis change is the unitary of the QW.*

The protocol consists of N iterations of the following steps:

1. *Alice picks a random bit* $w_A \in \{0, 1\}$ *and a value* $i_A \in \{0, 1, \ldots, 2P - 1\}$.
   - *If* $w_A = 0$*: Alice will prepare and send to Bob the* $2P$*-dimensional state* $|\psi_i\rangle = |i_A\rangle$.
   - *If* $w_A = 1$*: Alice will prepare and send to bob the* $2P$*-dimensional state* $|\psi_i\rangle = U_k^t(I_p \otimes F)\,|i_A\rangle$.
2. *Bob picks a random bit* $w_B \in \{0, 1\}$.
   - *If* $w_B = 0$*: Bob measures the received* $2P$ *dimensional state in the computational Z basis resulting in outcome* j_B.
   - *If* $w_B = 1$*: Bob measures in the QW basis (alternatively, he inverts the QW by applying* $\left(U_k^t\right)^{-1}$ *and measures the resulting state in the Z basis). The result is translated, in the obvious way, into an integer* j_B.

   Note that he measures both the position and coin, as opposite to the previous protocol, where the measurement for the key was only on the positions space.
3. *Alice and Bob reveal, via the authenticated classical channel, their choice of* w_A *and* w_B*. If* $w_A = w_B$*, they will use their values* i_A *and* j_B *to contribute toward their raw key. Otherwise, if* $w_A \neq w_B$*, they will discard this iteration.*

After the above process, Alice and Bob will use a cut-and-choose technique similar to Yao's [115], to check eavesdropping by choosing a suitable subset of non-discarded iterations for parameter estimation in the usual manner (discarding those chosen iterations from the raw key). This allows them to estimate the disturbance Q_Z *and* Q_W *in the Z and QW bases, respectively (i.e., in the absence of noise* $Q_Z = Q_W = 0$*).*

If this disturbance is "sufficiently low" (to be discussed below), the users proceed with error correction and privacy amplification in the usual manner.

4.1 Security

In order to prove the security of Protocol 2, we will construct, in the usual way, an equivalent entanglement-based protocol [11,62]. Proving security of this entanglement-based protocol will show the security of the prepare-and-measure version. This equivalence between entanglement-based and prepare-and-measure QKD protocols was first established by Bennett, Brassard and Mermin in [11]. Since then, the relationship between the presence of entanglement (and specifically the ability of the involved parties to certify or distill entanglement) and the security of prepare-and-measure QKD protocols has been developed and thoroughly investigated [27,28]. In this context, a quite common technique, when it comes to proving security of prepare-and-measure QKD protocols (such as the ones we present in this work), is to consider an equivalent entanglement-based protocol and prove its security [11,62]. We should stress that it can also hold even in the case that the devices are not trusted (device-independent QKD), given that some specific assumptions about the devices are made [113].

For this entanglement-based version, for each one of the N iterations, we make changes to steps (1) and (2), replacing them as follows:

New Step (1): Alice prepares the entangled state:

$$|\phi_0\rangle = \frac{1}{\sqrt{2P}} \sum_{i=0}^{2P-1} |i, i\rangle_{AB}$$

which lives in the $4P^2$ dimensional Hilbert space: $\left(\mathcal{H}_p \otimes \mathcal{H}_c\right)^{\otimes 2}$. She sends the second half (the Bob portion of $|\phi_0\rangle$) to Bob while keeping the first half (the Alice portion) in her private lab.

New Step (2): Alice and Bob choose independently two random bits w_A and w_B. If $w_A = 0$, Alice will measure her half of the entangled state in the computational Z basis; otherwise she will measure her half in the QW basis. Similarly for Bob and w_B. Let their measurement results in values be i_A on Alice's side and j_B on bob's side.

We now show the security of this entanglement-based version of the protocol. In the following proof, we will initially make three assumptions:

A1: Alice and Bob only use those iterations where $w_A = w_B = 0$ for their raw key.
A2: Eve is restricted to collective attacks (those whereby she attacks each iteration of the protocol independently and identically, but is free to perform a joint measurement of her ancilla at any future time of her choosing).
A3: Eve is the party that actually prepares the states which Alice and Bob hold.

Assumption A1 is made only to simplify the computation and may be discarded later (alternatively, one may bias the basis choice so that w_A and w_B are chosen to be 0 with high probability, thus increasing the efficiency of the protocol as is done for instance for BB84 in [63]). Assumption A2 may be removed later using a

de Finetti-type argument [26,86,87] (in this paper, we are only concerned with the asymptotic scenario, so the key-rate expression we derive will not be degraded). Note that removing A2 gives us the security. Assumption A3 gives greater advantage to the adversary; if we prove security using A3, then the "real-world" case, where assumption A3 is not used, will certainly be just as secure, if not even more.

In light of A2 and A3, Alice, Bob, and Eve, after N iterations of the protocol, hold a quantum state $\rho_{ABE}^{\otimes N}$, where $\rho_{ABE} \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E$ with $\mathcal{H}_A \equiv \mathcal{H}_B \equiv \mathcal{H}_p \otimes \mathcal{H}_c$. We should note that, since we consider Eve to be an all-powerful adversary, there are no restrictions on the specific form of her Hilbert space $\mathcal{H}_E$. Following error correction and privacy amplification, Alice and Bob will hold a secret key of size $\ell(N)$. Under the assumption of collective attacks (A2), we may use the Devetak–Winter key-rate expression [30] to compute:

$$r = \lim_{N\to\infty} \frac{\ell(N)}{N} = S(A|E) - H(A|B).$$

Let A_Z and A_W be the random variables describing Alice's system, when she measures in the Z or QW basis, respectively. Similarly, define B_Z and B_W. Under assumption A1, we are actually interested in the value:

$$r = S(A_Z|E) - H(A_Z|B_Z).$$

Computing $H(A_Z|B_Z)$ is trivial, given the observable probabilities:

$$p_{i,j}^{Z} = Pr(i_A = i \text{ and } j_B = j \mid w_A = w_B = 0). \tag{6}$$

The challenge is to determine a bound on the von Neumann entropy $S(A_Z|E)$.

To do so, we will use an uncertainty relation, proven in [12], which states that for any density operator σ_{ABE} acting on Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E$, if Alice and Bob make measurements using POVMs $\mathcal{M}_0 = \left\{M_x^{(0)}\right\}_x$ or $\mathcal{M}_1 = \left\{M_x^{(1)}\right\}_x$, then

$$S(A_0|E) + H(A_1|B) \geq \log \frac{1}{c}, \tag{7}$$

where

$$c = \max_{x,y} \left|\left| M_x^{(0)} M_y^{(1)} \right|\right|_\infty^2 \tag{8}$$

where we take $||\cdot||_\infty$ to be the operator norm and A_i to be the random variable describing Alice's system after measuring $\mathcal{M}_i$ (we will later, similarly, define B_i). Assuming measurements $\mathcal{M}_0$ are used for key distillation, simple algebra, as discussed in [12], yields the Devetak-Winter key-rate:

$$\begin{aligned} r = S(A_0|E) - H(A_0|B_0) &\geq \log \frac{1}{c} - H(A_0|B_0) - H(A_1|B) \\ &\geq \log \frac{1}{c} - H(A_0|B_0) - H(A_1|B_1). \end{aligned}$$

The last inequality follows from the basic fact that measurements can only increase entropy.

In our case, we have $M_x^{(0)} = |x\rangle\langle x|$ and $M_x^{(1)} = |\psi_x\rangle\langle\psi_x|$ for $x \in \{0, 1, \ldots, 2P-1\}$. Let $|\psi_x\rangle = \sum_{i=0}^{2P-1} \alpha_{x,i} |i\rangle$; then, it is easy to see that for all x, y

$$\left|\left| M_x^{(0)} M_y^{(1)} \right|\right|_\infty^2 = |\alpha_{y,x}|^2,$$

and therefore

$$c = \max_{x,y} |\alpha_{x,y}|^2, \tag{9}$$

a quantity which depends exclusively on the choice of the QW parameters and not on the noise in the channel. Therefore, Alice and Bob should choose optimal t, θ_k and P in order to minimize c (thereby maximizing the key-rate equation).

As we show in the next section, this analysis is sufficient to derive good key-rate bounds.

4.2 Evaluation

As mentioned above, the value of c depends solely on the QW parameters which are under Alice and Bob's control; therefore, it is to their advantage to choose a QW which minimizes this value (i.e., such that, after evolving for t steps, the probability of finding the walker at any particular position is small). It is easy to see that, as $t \to \infty$, the values $|\alpha_{x,y}|$ do not converge to a steady state which is why, usually, one considers the time-averaged distribution when analyzing QWs on the cycle [1,47]. However, in our QKD protocol, we do not care what happens at large t; instead, we wish to find an optimal t and one that is preferably not "too large" (the larger it is, the longer, in general, it might take Alice to prepare the state and Bob to reverse it). Notice that, while in the previous two-way protocol the choice of large t increased the security of the protocol, here the role of t is different. Larger t does not mean that it is harder for Eve to distinguish. It is just a parameter that Alice and Bob can tune to get optimal noise resistance, and in this context, small t (even $t = 1$ or 2) can lead to secure systems, as well. Different values of t correspond to different noise tolerances and our investigation is looking for optimal values of t where "optimal" means the highest theoretical noise tolerance for a *given* walk setting. If our system were to be implemented in practice, "optimal" would probably take a much different form and Alice and Bob would have to consider their device imperfections, etc. Such an investigation, though very interesting, is outside the scope of this paper, but it presents a relevant direction of future work.

We begin by looking at various walk parameters and finding the minimal value of c when $F = I_c$, the identity operator. Note that, on the circle, it makes sense only to consider odd P as even P would force the support of the probability amplitudes onto even or odd numbered nodes only thereby increasing the overall value of $|\alpha_{x,y}|$. We wrote a computer program to simulate the walk for time steps $t = 1, 2, \ldots, T_{\max}$ (for user-specified value $T_{\max}$) searching for the optimal value of t (i.e., a value for t

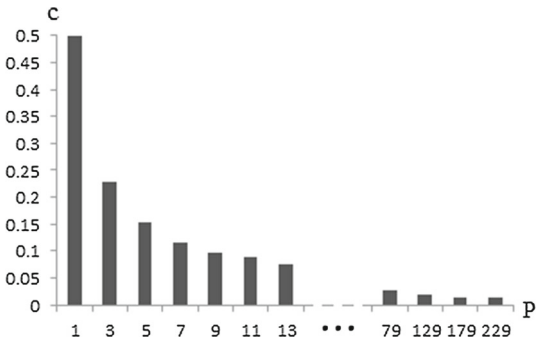


Fig. 3 Showing minimal value of c found by our program for given position space dimension P when $\theta = \pi/4$, $\phi = 0$ and $F = I_c$. When $P \leq 13$ we set $T_{\max} = 5000$; when $P \geq 79$ we set $T_{\max} = 50{,}000$. Note that, the smaller c is, the better for Alice and Bob. Note also that P is the dimension of the position space, *not* the number of qubits sent which would actually be $\lceil \log P \rceil + 1$ (where the extra "+1" is due to the coin)

whereby c is minimum). For the evaluation, we used a more general form of the coin rotation operator:

$$R_c(\theta, \phi) = \begin{pmatrix} e^{i\phi}\cos(\theta) & e^{i\phi}\sin(\theta) \\ -e^{-i\phi}\sin(\theta) & e^{-i\phi}\cos(\theta) \end{pmatrix},$$

The results for $\theta = \pi/4$, $\phi = 0$, and for various P are shown in Fig. 3.

Now that we can find the optimal choice of QW parameters for particular values of P and, more importantly for our work here, the resulting value of c. To this end, we have to compute our bound r and determine for what noise levels we can have $r > 0$. In practice, one would observe values $p^Z_{i,j}$ and $p^W_{i,j}$ [see Eq. (6) and define $p^W_{i,j}$ analogously] and use these to directly compute $H(A_Z|B_Z)$ and $H(A_W|B_W)$ as required by the key-rate equation. For the purpose of illustration in this paper, however, we will evaluate our key-rate bound assuming a generalized Pauli channel as discussed in [5] (see, in particular, Sect. 7 of that source). This channel maps an input state ρ to an output state $\mathcal{E}(\rho)$ defined as:

$$\mathcal{E}(\rho) = \sum_{m=0}^{2P-1} \sum_{n=0}^{2P-1} p_{m,n}\, \mathcal{U}_{m,n}\, \rho\, \mathcal{U}^*_{m,n}, \tag{10}$$

where

$$\mathcal{U}_{m,n} = \sum_{k=0}^{2P-1} e^{\pi \cdot i \cdot k \cdot n/P}\, |k+m\rangle\langle k|\,. \tag{11}$$

That is, this channel $\mathcal{E}(\cdot)$ models an adversary's attack which induces phase and flip errors with probabilities denoted by $p_{m,n}$. In our numerical computations to follow, we will use:

$$p_{i,j} = \begin{cases} 1 - E_r & \text{if } i = j = 0 \\ \dfrac{E_r}{(2P)^2 - 1} & \text{otherwise} \end{cases}. \tag{12}$$

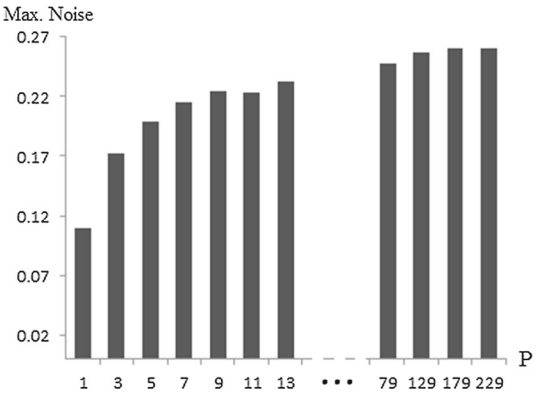


Fig. 4 Showing the maximally tolerated noise level for our protocol using parameters found in Fig. 3 and using the quantum channel described by Eqs. (10) and (12). The lack of increase in noise tolerance from $P = 9$ to $P = 11$ (while other choices caused an increase) indicates that $T_{\max}$ was too low. Note that, when $P = 1$, we recover the BB84 tolerance of $Q = 0.11$ as expected. Also note that, when $P = 229$, the maximal tolerated noise is $Q = 0.261$

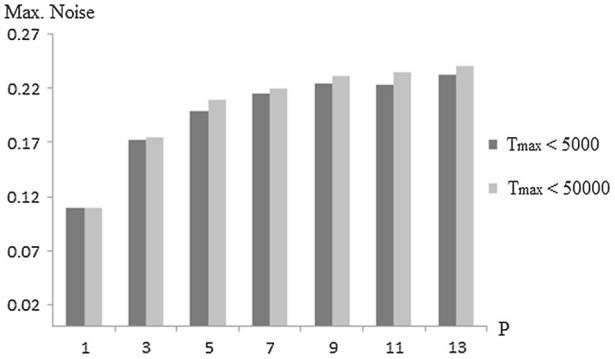


Fig. 5 Comparing the maximally tolerated noise when t is allowed to be as large as 50,000 (light gray) or only 5000 (dark gray); again when $F = I$ and $\phi = 0$. In this case, when $P = 13$ and $T_{\max} = 50{,}000$, the maximal tolerated noise Q is $Q = 0.241$

It is clear that $\sum_{i,j} p_{i,j} = 1$. Furthermore, when $E_r = 0$, we have $\sum_i p^Z_{i,i} = \sum_i p^W_{i,i} = 1$ (i.e., there is no disturbance in the channel), while as E_r increases, the disturbance also increases.

Finally, we define the total noise in the channel to be:

$$Q = \sum_{a \neq b} p^Z_{a,b} = \sum_{a \neq b} Pr\big(A_Z = a \text{ and } B_Z = b \mid w_A = w_B = 0\big).$$

That is to say, Q represents the quantum error rate (QER) of the channel.

The maximally tolerated QER, for those QWs analyzed in Fig. 3, and using the above described noise model, is shown in Fig. 4. Note that, when $P = 1$ and $t = 1$, we recover the BB84 limit of 11% which is to be expected since, with these choice of parameters, we are essentially running the BB84 protocol.

Observe in Fig. 4 that there is a lack of increase when $P = 9$ and $P = 11$; this indicates that our choice of $T_{\max} = 5000$ was too low. Running our simulator again with $T_{\max} = 50000$ for these small P values yields a maximally tolerated noise level shown in Fig. 5.

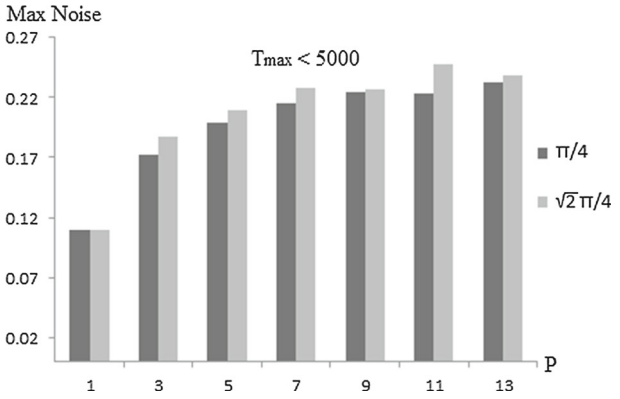


Fig. 6 Comparing the maximal tolerated noise levels of the QKD protocol when $\theta = \pi/4$ (dark gray) and $\theta = \sqrt{2}\pi/4$ (light gray). In this chart, $T_{\max} = 5000$ which, observing the "drop" in tolerated noise when P goes from 11 to 13, is too small. See also Fig. 7 for the same chart when $T_{\max} = 50{,}000$

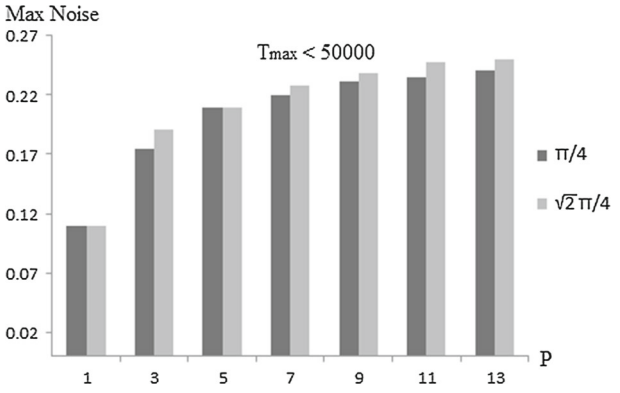


Fig. 7 Comparing the maximal tolerated noise levels of the QKD protocol when $\theta = \pi/4$ (dark gray) and $\theta = \sqrt{2}\pi/4$ (light gray). In this chart, $T_{\max} = 50{,}000$. In all cases, the QW parameter $\theta = \sqrt{2}\pi/4$ produces a more secure QKD protocol for this upper-bound on t. Note that, as $T_{\max} \to \infty$, they may produce equally secure protocols; this, as discussed in the text, is an open question. In this case, when $P = 13$ and $\theta = \sqrt{2}\pi/4$, the maximally tolerated noise is 0.25 (compared to 0.241 when $\theta = \pi/4$)

Finally, we re-run the simulator, using $T_{\max} = 5000$ and $T_{\max} = 50000$ for a different QW parameter of $\theta = \sqrt{2}\pi/4$ which, for these particular upper-bounds on t yield a higher tolerated noise as shown in Figs. 6 and 7. We comment that, if $T_{\max}$ were larger, the two QWs may produce a QKD protocol with the same tolerated noise; however, for these "smaller" bounds on t, the QW with parameter $\theta = \sqrt{2}\pi/4$ produces a more secure protocol than when $\theta = \pi/4$. Since smaller t implies a more efficient protocol, this is an advantage. This opens two very interesting questions: first, do these QWs produce equivalent noise tolerances as $T_{\max} \to \infty$? Second, what other values of θ produce even more secure QKD protocols for small $T_{\max}$? We comment that we also ran this numerical experiment for $\theta = \pi/5$ and $\theta = \pi/3$ but got worse noise tolerances.

From the above, it is clear that careful choice of the QW parameters is vital for producing a QKD protocol tolerant of high noise channels. To investigate this further,

we simulate the QW for all $\theta, \phi \in \{k\pi/10 \mid k = 0, 1, \ldots, 10\}$. Furthermore, for each setting, we also consider the use of $F = I$, $F = X$, and $F = Y$, where:

$$X = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad Y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}.$$

For each setting, we find the optimal choice of time $t \leq 5000$ which produces a minimal c. We then take this value and determine the highest disturbance the resulting protocol can withstand. The respective data are summarized in Table 1.

Note that, for some data points (e.g., when $P = 5$ and $F = I$), there is a drop in the maximum tolerated noise. This is a consequence either of setting $T_{\max}$ too small, or we need to simulate more QW parameters (as is done in Table 2). For example, when we set $T_{\max} = 50000$, for $P = 5$ and $F = I$, we get a maximum noise tolerance of 0.236 when $t = 40847$. Note also, that setting $F = Y$ achieves the best result for this test, $Q_{\max} = 0.284$.

In Table 2, we carried out the same experiment; however, this time searching over QW parameters in the set $\theta, \phi \in \{k\pi/20 \mid k = 0, 1, \ldots, 20\}$. Again, the best result for this case is $Q_{\max} = 0.284$ and is achieved when considering $F = Y$.

As mentioned at the beginning of this Section, all the numerical results were obtained by simulating the evolution of the QW on a custom QW simulator that we wrote. However, we also verified the results through an alternative technique, namely by computing the probability amplitudes of the QW using the standard Fourier method (see, e.g., [4,105]) of analyzing QWs. The results obtained by both methods agree with each other.

We would also like to stress that, while the use of high-dimensional walks is "ideal" from a noise-tolerance perspective, this is not required—indeed, even with very small dimensions our protocol can tolerate the same level of noise as BB84. In particular, notice in Fig. 4, that already for the minimum $P = 1$, we obtain the BB84 noise tolerance, which increases further for relatively small dimensions $P = 3$, $P = 5$ and so on. We also simulated higher dimensions to investigate the theoretical properties of our systems for different walks. However, such high dimensions are not required to get a robust protocol. Therefore, while the practical implementation of a high-dimensional protocol might be quite complex, our protocol is robust even for relatively small dimensions; thus, the experimental challenges in a potential practical implementation could be reduced. Moreover, we should emphasize that recently there has been a remarkable progress in the experimental generation and manipulation of high-dimensional entangled states. In particular, in [29,46,54,76,107], the experimental generation of bipartite high-dimensional entangled states has been demonstrated, while in [6,44,58,78] there have been proposed experimental techniques for performing measurements in such states. Furthermore, the generation of multi-partite high-dimensional entangled states has been reported [35,42,74], as well as the generation of big arrays of such states [55].

Finally, we note that the protocol's security is not compromised by considering the existence or not of quantum memories. It is sufficient to consider the prepare-and-measure form of the protocol. Eve needs a quantum memory to perform her attack, as she needs to save her ancillary system throughout the execution of the protocol. In the contrary, the secure key distribution between Alice and Bob does not require

Table 1 Showing the optimal choice of QW parameters to maximize the noise tolerance ($Q_{\max}$) of the resulting protocol

P	$F = I$					$F = X$					$F = Y$				
	θ	ϕ	t	c	$Q_{\max}$	θ	ϕ	t	c	$Q_{\max}$	θ	ϕ	t	c	$Q_{\max}$
3	0.4π	0.2π	4584	0.171	0.220	0.8π	0.8π	3994	0.181	0.211	0.7π	0	1502	0.167	0.225
5	0.7π	π	4340	0.147	0.205	0.9π	0.5π	3870	0.132	0.22	0.3π	0	3748	0.106	0.253
7	0.6π	0.9π	3946	0.088	0.252	0.7π	0.8π	3391	0.099	0.236	0.3π	0.5π	1275	0.083	0.261
9	0.6π	0.6π	1269	0.077	0.252	0.9π	0.7π	3041	0.079	0.250	0.3π	0.5π	965	0.069	0.267
11	0.6π	0.4π	1221	0.069	0.252	0.8π	0.4π	481	0.0724	0.245	0.7π	0.5π	277	0.054	0.284

For this data, we searched for QWs with at most $T_{\max} = 5000$ steps and with parameters $\theta, \phi \in \{k\pi/10 \mid k = 0, 1, \ldots, 10\}$

Table 2 Showing the optimal choice of QW parameters to maximize the noise tolerance ($Q_{\max}$) of the resulting protocol

P	$F = I$					$F = X$					$F = Y$				
	θ	ϕ	t	c	$Q_{\max}$	θ	ϕ	t	c	$Q_{\max}$	θ	ϕ	t	c	$Q_{\max}$
3	0.4π	0.2π	4584	0.171	0.220	0.25π	0.15π	2402	0.173	0.218	0.7π	0	1502	0.167	0.225
5	0.6π	0.85π	3258	0.116	0.239	0.8π	0.25π	4659	0.124	0.229	0.3π	0	3748	0.106	0.253
7	0.6π	0.9π	3946	0.088	0.252	0.05π	0.95π	3739	0.091	0.248	0.35π	0.5π	517	0.081	0.265
9	0.45π	0.95π	2531	0.075	0.257	0.85π	0.05π	1669	0.078	0.251	0.45π	0.5π	1240	0.064	0.276
11	0.55π	0.75π	1826	0.059	0.272	0.25π	0.25π	2223	0.069	0.252	0.7π	0.5π	277	0.054	0.284

For this data, we searched for QWs with at most $T_{\max} = 5000$ steps and with parameters $\theta, \phi \in \{k\pi/20 \mid k = 0, 1, \ldots, 20\}$

any quantum memory. Therefore, if Eve does not have a quantum memory she cannot attack, while if even she has one and attacks, Alice and bob can defend against it and securely share a key at the end. Notice, that even if we consider the entanglement-based version of the protocol, again the security is independent of any quantum memory requirements, as Eve for her attack needs a more stable quantum memory than Alice and Bob need to defend against it and securely distill the key.

5 Semi-quantum key distribution scheme

In this section, we will present a SQKD protocol based on QWs. Semi-quantum protocols can be seen as practical instances of QKD, since they involve less quantum hardware, as one of the parties is completely classical. Also, they are interesting from a theoretical point of view, as they can be treated as a measure of the "quantumness" needed for a protocol to surpass the security of its classical counterparts. It is assumed that one of the parties, e.g., Alice, is fully quantum, i.e., possesses quantum equipment, while the other party, e.g., Bob, is restricted to classical operations only, and for that reason is usually called the *classical party*.

Semi-quantum protocols rely on a two-way quantum channel allowing a quantum state to travel from Alice to Bob, then back to Alice. When first introduced by Boyer et al. in [14], these classical operations involved Bob either measuring the incoming qubit in the $Z = \{|0\rangle, |1\rangle\}$ basis, or reflecting the incoming qubit, bouncing it back to Alice undisturbed. For our purposes, we extend this definition of "classical" operations to operate with higher dimensional systems. As we do not want to restrict ourselves necessarily to qubit encodings (and thus, dimensions that are powers of two), we will say that Bob, on receipt of an D dimensional quantum state $|\psi\rangle$, may choose to do one of two operations:

1. Measure and Resend: Bob may subject the D-dimensional quantum state to a measurement in the computational basis spanned by states: $\{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$. He will then prepare a new D-dimensional quantum state in this same computational basis based on the result of his measurement. Namely, if he observes $|r\rangle$ for $r \in \{0, 1, \ldots, D-1\}$, he will send to Alice the quantum state $|r\rangle$.
2. Reflect: Bob may ignore the incoming D-dimensional quantum state and reflect it back to Alice. In this case, he learns nothing about its state.

With these restrictions on the part of the classical user defined, we now depict and describe our protocol (Fig. 8):

Protocol 3 *Semi-quantum key distribution scheme*

Inputs for the protocol

- $|l, s\rangle$, *the initial state of the QW, where* $l \in \mathcal{L} = \{0, \ldots, P-1\}$ *is the initial position of the walker, and* $s \in \mathcal{S} = \{R, L\}$ *gives the initial coin state.*
- $U_{walk} = (U_k)^t \in \mathcal{Q}$, *the evolution of the QW, where* $k \in \mathcal{K} = \{1, 2, \ldots, K\}$ *is the choice of a single step unitary* U_k, *and* $t \in \mathcal{T} = \{T_0, \ldots, T_{max}\}$ *is the number of steps of the QW. Thus,* $\mathcal{Q}$ *is the set of all possible QWs. Note that* $\mathcal{Q}$ *is publicly known.*

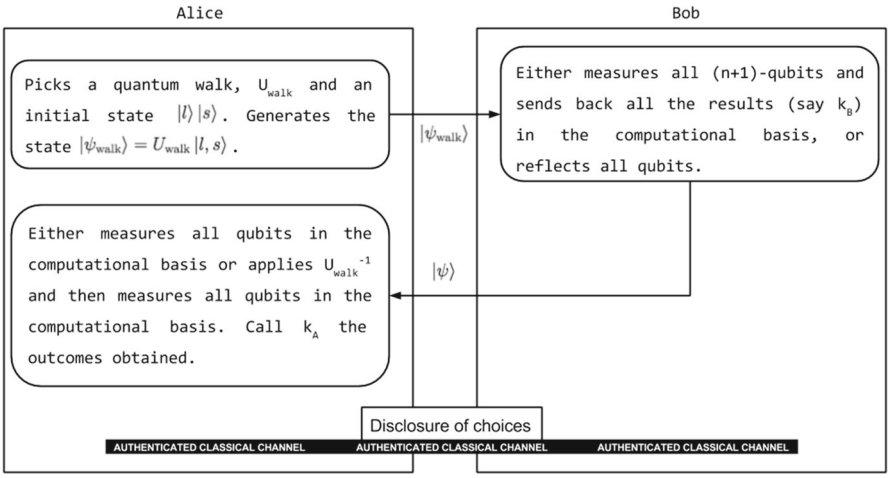


Fig. 8 Description of the basic steps of Protocol 3

Quantum state generation

– *Alice chooses uniformly at random* $l \in \mathcal{L} = \{0, 1, \ldots, P-1\}$ *and* $s \in \mathcal{S} = \{R, L\}$. *She also chooses a random QW operator* $U_{walk} \in \mathcal{Q}$ *according to a publicly known distribution (e.g., uniform). She then prepares the following state:*

$$|\psi_{walk}\rangle = U_{walk}\,|l, s\rangle\,.$$

– *Alice sends this state to Bob.*

Classical operations by Bob

Bob chooses either to measure-and-resend the quantum state in the computational basis $\{|0\rangle, |1\rangle, \ldots, |2P-1\rangle\}$ *(note that, in this protocol as well, he measures both the position and coin in order to obtain the key, thus his measurement, and subsequent preparation, is of dimension* $2P$*); or he will reflect the quantum state back to Alice.*

Alice's final step

Alice chooses one of the following two options:

– *She measures the returning quantum state in the computational basis and saves the result as* κ_A.
– *She first applies the inverse QW,* U_{walk}^{-1}, *and then measures in the computational basis. Note that, in the absence of noise, if Bob reflects, her measurement outcome should be* $|l, s\rangle$.

Disclosure

Alice discloses her choice of operation and Bob discloses his choice either to measure and resend or reflect.

Iterations
The above process is repeated N times.

Results

- *Every time bob measures and resends and Alice measures in the computational basis, the parties add $1 + \log P$ bits to their final raw key. After N iterations, the raw key will consist of $N(1 + \log P)/4$ bits, on average. Considering that Alice and Bob choose independently, uniformly and at random their actions at each iteration, then—on average—$N/4$ iterations will contribute to the raw key.*
- *Every time Bob reflects and Alice measures after applying the inverse QW, the outcome of her measurement (l_m, s_m) should be what she initially used to generate the QW state (i.e., it should be that $l = l_m$ and $s = s_m$). These iterations, together with some randomly chosen iterations of the first type (where Bob measures and resends), are used for error detection.*
- *The other iterations are discarded.*

5.1 Proof of robustness

As with the first protocol we proposed in this paper, the reliance on a two-way quantum channel greatly complicates the security analysis. It was only recently that several SQKD protocols were proven secure [51–53,117]. However, the proof techniques developed in those works assumed qubit-level systems. In our case, not only must we contend with a two-way channel, but also with the fact that the quantum states traveling between Alice and Bob are of dimensions higher than 2. This leads to significant challenges in any security analysis.

Therefore, to analyze the security of this protocol, we will prove that it is *robust* as defined in [13,14]. That is, for any attack which Eve may perform which causes her to gain information on the raw key, this attack must necessarily lead to a disturbance in the channel which can be detected with non-zero probability by Alice and Bob. Typically, proving robustness is a first step in the security analysis of semi-quantum cryptographic protocols.

Theorem 4 *If $I \in \mathcal{Q}$ (where I is the identity operator on the joint $2P$ dimensional system) and if, for every $(l, s), (l', s') \in \{0, 1, \ldots, P-1\} \times \{R, L\}$ there exists a $U_{walk} \in \mathcal{Q}$ and initial state $|l_0, s_0\rangle$ (all possibly depending on the choice of (l, s) and (l', s')) such that $\langle l, s|U_{walk}|l_0, s_0\rangle \neq 0$ and $\langle l', s'|U_{walk}|l_0, s_0\rangle \neq 0$, then the SQKD protocol based on QWs is robust.*

Proof We will assume, similarly to [49,119], that Alice sends each (in our case $2P$-dimensional) quantum state, only after she receives one from Bob (excepting, of course, the first iteration). In this case, Eve's most general attack consists of a collection of unitary operators $\left\{(U_F^{(i)}, U_R^{(i)})\right\}_{i=1}^{N}$ where, on iteration i of the protocol, she applies $U_F^{(i)}$ in the forward channel (as the quantum state travels from Alice to Bob) and $U_R^{(i)}$ in the reverse channel. These operators act on the $2P$-dimensional quantum

state and Eve's private quantum memory. We make no assumptions about how these operators are chosen—for instance, Eve may choose them "on the fly"; that is, she may choose operator $U_F^{(2)}$ after attacking with $U_F^{(1)}$.

Consider the first iteration $i = 1$. We assume, without loss of generality, that Eve's quantum memory is cleared to some pure "zero" state, denoted by $|\chi\rangle$, known to her.

In the remainder of this proof, we will treat the position space and the coin space as a single space Σ of dimension $2P$.

We may describe the action of $U_F^{(1)}$ on basis states as follows

$$U_F^{(1)} |i, \chi\rangle = \sum_{j=0}^{2P-1} |j, e_i^j\rangle ,$$

where $|e_i^j\rangle$ are arbitrary states in Eve's ancillary system. These states are not necessarily normalized nor orthogonal; the unitarity of $U_F^{(1)}$ imposes some restrictions on them which we will use later.

With non-zero probability, this iteration may be used for error detection. It is also possible that Alice chose to use $I \in \mathcal{Q}$ in this iteration and, thus, she sends the quantum state $|\sigma\rangle$ to Bob, for $\sigma \in \Sigma$. Furthermore, Bob chooses to measure and resend with non-zero probability. Therefore, to avoid detection, it must be that $|e_i^j\rangle \equiv 0$ for all $i \neq j$, and the unitarity of $U_F^{(1)}$ yields $\langle e_i^i | e_i^i \rangle = 1$ for all i. Thus:

$$U_F^{(1)} |i, \chi\rangle = |i, e_i^i\rangle , \forall i = 0, 1, \ldots, 2P - 1.$$

Now, consider $U_R^{(1)}$, the attack applied in the reverse channel. We may write its action as follows:

$$U_R^{(1)} |i, e_i^i\rangle = \sum_{w=0}^{2P-1} |w, e_{i,i}^w\rangle .$$

The same argument as before applies: in particular, with non-zero probability Alice and Bob will use this iteration to check for errors, and so it must be that $|e_{i,i}^w\rangle \equiv 0$ for $i \neq w$. Thus

$$U_R^{(1)} |i, e_i^i\rangle = |i, e_{i,i}^i\rangle = |i, f_i\rangle , \forall i = 0, 1, \ldots, 2P - 1,$$

where we defined $|f_i\rangle \equiv |e_{i,i}^i\rangle$ for ease of notation.

Now, assume that Alice chooses a QW operator $U_{\text{walk}} \in \mathcal{Q}$, with $U_{\text{walk}} \neq I$. Let $|\sigma\rangle$ be the initial state she prepares (σ chosen at random from Σ). In this case, the quantum state she sends to Bob may be written as:

$$U_{\text{walk}} |\sigma\rangle = |\psi_\sigma\rangle = \sum_{i=0}^{2P-1} \alpha_i |i\rangle .$$

Assume that U_{walk} is chosen so that at least two of the α_i's are non-zero (such QWs exist by hypothesis). If Bob reflects, the qubit state arriving at Alice's lab, after Eve's attack on both channels, is

$$U_R^{(1)} U_F^{(1)} (U_{\text{walk}} \otimes I_E) \, |\sigma, \chi\rangle = \sum_i \alpha_i \, |i, f_i\rangle \,, \tag{13}$$

where I_E is the identity operator on Eve's ancilla.

Alice will subsequently apply the inverse QW operator and measure the resulting state, expecting to find $|\sigma\rangle$. This is equivalent to her measuring in the QW basis $\{|\psi_0\rangle \,, |\psi_1\rangle \,, \ldots \,, |\psi_{2P-1}\rangle\}$, where $|\psi_i\rangle = U_{\text{walk}} \, |i\rangle$, and expecting to observe $|\psi_\sigma\rangle$. In this QW basis, we clearly have

$$|i\rangle = \sum_{j=0}^{2P-1} \langle \psi_j | i \rangle \, |\psi_j\rangle \,,$$

from which, we may write Eq. (13) as:

$$\sum_{i=0}^{2P-1} \alpha_i \left(\sum_{j=0}^{2P-1} \langle \psi_j | i \rangle \, |\psi_j\rangle \right) \otimes |f_i\rangle \tag{14}$$

$$= \sum_{j=0}^{2P-1} |\psi_j\rangle \otimes \left(\sum_{i=0}^{2P-1} \alpha_i \, \langle \psi_j | i \rangle \, |f_i\rangle \right). \tag{15}$$

Let p be the probability that this iteration does not result in an error—i.e., the probability that Alice measures $|\psi_\sigma\rangle$. From the above equation:

$$p = \left| \sum_{i=0}^{2P-1} \alpha_i \, \langle \psi_\sigma | i \rangle \, |f_i\rangle \right|^2 .$$

Noticing that $\langle \psi_\sigma | i \rangle = \alpha_i^*$ (since $|\psi_\sigma\rangle = \sum_i \alpha_i \, |i\rangle$), and also $\langle f_i | f_i \rangle = 1$ (due to the unitarity of $U_R^{(1)}$), we find:

$$p = \left| \sum_i |\alpha_i|^2 \, |f_i\rangle \right|^2 = \sum_i |\alpha_i|^4 + 2 \sum_{i>j\geq 0} |\alpha_i|^2 |\alpha_j|^2 \text{Re}(\langle f_i | f_j \rangle).$$

When $|f_i\rangle \equiv |f_j\rangle = |F\rangle$, for all i, j, the above quantity attains its maximum of $p = 1$. In this case, after Eve's attack, the system described by Eq. (13) is $\sum_i \alpha_i \, |i\rangle \otimes |F\rangle = |\psi_\sigma\rangle \otimes |F\rangle$. Due to the Cauchy-Schwarz inequality $\text{Re}(\langle f_i | f_j \rangle) \leq 1$. If, however, one or more of the $\text{Re}(\langle f_i | f_j \rangle) < 1$ for any of the $(|f_i\rangle \,, |f_j\rangle)$ pairs which

appear in the expression above (i.e., for those where α_i and α_j are non-zero), it is obvious that $p < 1$ and so Eve would be detected.

Therefore, to avoid detection, it must be that $\mathrm{Re}(\langle f_i|f_j\rangle) = 1$ for all i, j where α_i and α_j are non-zero, implying $|f_i\rangle \equiv |f_j\rangle$. Indeed, if we write $|f_j\rangle = x\,|f_i\rangle + y\,|\zeta\rangle$, where $\langle f_i|\zeta\rangle = 0$, then $\mathrm{Re}(\langle f_i|f_j\rangle) = 1 = \mathrm{Re}(x)$. Of course $|x|^2 + |y|^2 = 1$ (since $\langle f_j|f_j\rangle = 1$) and so:

$$|x|^2 + |y|^2 = 1 \Rightarrow \mathrm{Re}^2 x + \mathrm{Im}^2 x + |y|^2 = 1 \Rightarrow \mathrm{Im}^2 x + |y|^2 = 0.$$

This implies both $\mathrm{Im}(x) = 0$ and $y = 0$. Since $\mathrm{Re}(x) = 1$, we conclude $x = 1$ and so $|f_i\rangle = |f_j\rangle$.

Since Alice could have chosen any QW in $\mathcal{Q}$, all possible (i, j) pairs are covered (i.e., at least one QW in $\mathcal{Q}$ is guaranteed to produce a state where α_i and α_j are non-zero) and since Eve does not know which QW was chosen, it must be that $|f_i\rangle \equiv |f_j\rangle \equiv |F\rangle$ for all i, j.

Thus, after the first iteration, to avoid detection, it must be that the state of Eve's quantum memory is in the state $|F\rangle$, independently of Alice's and Bob's raw key and operations. Thus, Eve is not able to extract any information during the first iteration. Furthermore, since she is fully aware of the state of her quantum memory in this case (i.e., she knows the state $|F\rangle$), the above arguments may be repeated inductively for the remaining iterations of the protocol, leading to the conclusion that the protocol is robust. □

The above proof of robustness placed certain requirements on the set of QW $\mathcal{Q}$, but can such a set even exist? We show that, at least for all odd P, such a set may be easily constructed.

Lemma 1 *If P is odd, then there exists a set of QWs $\mathcal{Q}$ which satisfy the requirements of Theorem 4.*

Proof Let $(l, s), (l', s') \in \{0, 1, \ldots, P-1\} \times \{R, L\}$. We construct a QW $U_{l,s,l',s'}$ and an initial state $|l_0, s_0\rangle$ such that $\langle l, s|U_{l,s,l',s'}|l_0, s_0\rangle \neq 0$ and $\langle l', s'|U_{l,s,l',s'}|l_0, s_0\rangle \neq 0$.

Since P is odd, there exits a position index $q \in \{0, 1, \ldots, P-1\}$ and a value $q_0 \in \mathbb{Z}$ such that $|q_0| < P$, $q - q_0 \equiv l \pmod P$, and $q + q_0 \equiv l' \pmod P$. We assume that $q_0 \geq 0$; if $q_0 < 0$ the result is symmetric by simply "flipping" l with l' (in which case q_0 becomes non-negative).

The shift operator S for our QW is simply the usual

$$S = \sum_{i=0}^{P-1} |i-1\rangle\langle i| \otimes |R\rangle\langle R| + \sum_{i=0}^{P-1} |i+1\rangle\langle i| \otimes |L\rangle\langle L|,$$

where all arithmetic, of course, is done modulo P. Our coin operator will simply be the Hadamard coin:

$$R_c = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

We claim the desired operator is $U_{l,s,l',s'} = \left[(I_p \otimes R_c) \cdot S\right]^{t+1}$. (Note that the shift operator is applied before the coin in this case to simplify the construction) Now, consider the initial state $|q+1, R\rangle$. After the first step of the QW (i.e., after applying $(I_p \otimes R_c) \cdot S$), the QW evolves to the state $\frac{1}{\sqrt{2}} |q\rangle (|R\rangle + |L\rangle)$. It is not difficult to see that, after t additional steps with this QW, *but before the final application of* $I_p \otimes R_c$ *on the* $(t+1)$*-th step*, the quantum state evolves to:

$$\alpha |l, R\rangle + \beta |l', L\rangle + |\phi\rangle ,$$

where $|\alpha| \neq 0$, $|\beta| \neq 0$, and $|\phi\rangle$ is a non-normalized state orthogonal to both $|l, R\rangle$ and $|l', L\rangle$. Finally, after the last $I_p \otimes R_c$, the state becomes

$$\begin{aligned} U_{l,s,l',s'} |q+1, R\rangle = \frac{1}{\sqrt{2}} &(\alpha |l, R\rangle + \alpha |l, L\rangle \\ &+\beta |l', R\rangle - \beta |l', L\rangle) + |\phi'\rangle , \end{aligned}$$

with $|\phi'\rangle$ being a state orthogonal to $|l, R\rangle , |l, L\rangle , |l', R\rangle$, and $|l', L\rangle$, thus yielding the desired state. Taking $\mathcal{Q} = \bigcup_{l,s,l',s'} \left\{U_{l,s,l',s'}\right\} \cup \{I\}$ proves the result. □

Finally, we should notice that the robustness of this SQKD protocol is independent of the existence or absence of quantum memories. In fact, Eve's attack requires a stable quantum memory, in which she keeps her ancillary system during the execution of the protocol. On the other hand, Alice does not need any quantum memory in order to share the key with Bob at the end, and Bob is, of course, restricted to classical operations. Therefore, without a quantum memory Eve cannot even conduct the attack, whereas even if she has access to a quantum memory, she is not able to extract any useful information about the key without being detected by Alice and Bob.

6 Practical attacks

While our paper presents theoretical cryptographic proposals, and a detailed analysis of practical attacks is out of its scope, it is worthy presenting a short discussion of possible attacks and countermeasures for the case of optical implementations. The term practical attacks refers to attacks during which Eve is taking advantage of possible loopholes in the implementation of the protocols, i.e., the fact that the setups used for the implementation of the protocols are not perfect, can seriously compromise the security of the key. Such attacks have been thoroughly investigated in the literature, and several countermeasures have been proposed in different setups and scenarios. For an overview of the recent progress and current status of this area of QKD, see the following detailed reviews [9,31,32,45,95].

One of the most studied of such attacks is the photon number splitting attack (PNS), which is based on the fact that there are no perfect single-photon sources [15,69, 70]. Instead, the current sources emit in general multi-photon pulses, whose photon number statistics are described by a Poisson distribution. Eve, who is considered all powerful and bounded only by the laws of physics, can thus, by placing herself in

front of Alice, detect genuine multi-photon pulses, extract one photon from each, and send the rest to Bob through a lossless channel, while blocking single-photon pulses. Due to the fact that the quantum channel connecting Alice and Bob has losses exponential in the channel length, there exists a maximal distance, known to Eve, below which Bob is not able to spot Eve's interference. By storing the extracted photons in her quantum memory, Eve can measure them in the correct basis upon the classical communication between Alice and Bob, during which they publicly reveal their choices of preparation/measurement bases. Since all the photons of the same pulse are in the same state, Eve thus has the key shared by Alice and Bob.

The standard technique used to defend against a PNS attack is by introducing the so-called decoy states [43]. In addition to the signal states, from which the key is obtained, Alice sends coherent states $|e^{i\theta}|\alpha|\rangle$, with phase θ chosen uniformly at random, and the variable intensity $I \propto |\alpha|$. Note that such decoy pulses are to Eve indistinguishable from the signal ones. Thus, Alice and Bob can subsequently detect Eve's interference (extracting single photons from multi-photon pulses) by comparing the yields of signal and decoy states (given the channel loss ℓ, the yield y is defined as $y = 1 - \ell$ [43]). For more details, see [64], as well as subsequent improvements and modifications [41,68,72,91,109–111]. This method, developed for standard one-way QKD schemes, can be straightforwardly applied to our second proposal, which is a one-way protocol as well. It can also be applied to our first and third two-way proposals. Indeed, in our first proposal, as Bob does not perform any measurement, it is Alice who performs the yield estimation upon receiving back the pulses. The same can be done by Alice alone for the photons reflected by Bob in our third proposal, in which in addition the yield check could be done for the pulses measured by Bob. As mentioned above, the details of the techniques depend on particular implementations and are beyond the scope of our theoretical study.

While the PNS attack is applicable to most of the protocols that use imperfect photon sources, the above description of its particular implementation is given on the example of a standard QKD *one-way* scheme. Thus, it has to be re-examined when applied to different protocols. The crucial feature of the standard QKD protocol is the exchange of classical information between Alice and Bob, which allows Eve to extract the key exchanged. Therefore, since such exchange is present in our second and third protocol, the above described PNS attack is applicable to those protocols as well. Note though that in the case of the third, *two-way* protocol, Eve can possibly extract information only upon intercepting the pulses re-sent from Bob to Alice. Indeed, in the third protocol, the key is obtained from the cases in which Bob and, upon receiving them back, Alice too, perform measurements in the computational basis, thus sharing the same set of bits. Extracting photons from the pulse before it came to bob, and consequently before his measurement, gives Eve no information about the key.

Nevertheless, our first, *two-way* QKD protocol, is considerably different from the standard QKD ones, as Alice and Bob reveal *no classical information* regarding their quantum operations (they only exchange information regarding the cases used for verification procedure, which do not contribute to the key generation). Thus, Eve's task is more difficult than in the case of standard QKD protocols. What Eve can do is to extract *two* photons from each pulse, one on the way from Alice to Bob, and another on the way back to Alice, and compare their states, $|\psi\rangle$ and $|\psi(r)\rangle$, in the attempt to

learn the key r. Note that, even in the noiseless scenario, the described comparison does not have perfect efficiency, unlike the standard application of the PNS attack in which Eve learns the key with certainty. Moreover, in the case of our protocol, Eve can attack only three or more photon pulses, thus decreasing the efficiency of her attack with respect to the standard one, which makes use of more probable two-photon pulses as well. For example, for the commonly used order of the mean photons per pulse, $\mu = 0.2$, the probability for emitting three or more photons is $p(n \geq 3) \approx 0.001$, while the probability to emit exactly two photons (the "deficit" with respect to the standard PNS attack) is of the order of magnitude higher, $p(n = 2) \approx 0.016$, where n is the number of photons per pulse emitted.

Finally, we would like to note that, although in practical attacks Eve is assumed to be all powerful, exceeding the current technological equipment used by everyday users, not all practical attacks are based on the same level of subtle equipment. In the case of the PNS attack, Eve should be able to perform photon non-demolition number measurements, a task beyond any current and (at least mid-term) foreseeable technology.

Nevertheless, there exist other practical protocols that do not requite such sophisticated technology. Below, we briefly analyze three such kinds of attacks, extensively studied in the literature: the Trojan horse, the detector blinding and the time-shift attacks.

The Trojan horse attack is one of the first attacks ever considered and since then it has been thoroughly investigated and continuously developed in different contexts. In a nutshell, Trojan horse attacks benefits from the imperfections in the quantum channel between Alice and Bob that allows for Eve's interference by modulating Alice's pulses, sending them to Bob and analyzing the reflected/backscattered signal [67,92,104]. The first such attack benefitted from the detector imperfections, by collecting the light emitted upon the detection of the photons [56]. To counter such attacks, introducing simple optical isolators suffice in one-way protocols, while for two-way protocols one needs to introduce additional monitoring detectors [38].

Furthermore, we would like to briefly discuss two more attacks, namely the detector blinding and the time-shift attacks, which are both considered in the broader context of intercept and resend with faked states attacks [61]. In general, during an intercept and resend with faked states attack, Eve is not trying to extract information about the key from the original states that the legitimate parties exchange. Instead, she generates and sends to them classical or quantum light pulses, which are tailored in a way that she can control their measurement outcomes, while she is blocking the original states. At the end of such an attack, Eve and the legitimate parties share the same key, without Alice and Bob being able to detect Eve's interference. In both the aforementioned attacks, Eve is taking advantage of loopholes in the performance and efficiency of the detectors of the legitimate parties.

First, we consider the detector blinding attacks to standard one-way QKD protocols [71,112]. Eve first intercepts the state that Alice sends to Bob and measures it in one of the two possible bases, that she randomly chooses. Then, she sends to one of Bob's detectors a bright light pulse according to her measurement outcome. Note that the intensity of the bright light pulse is just a bit above the detector's threshold. If Bob chooses to measure in the same basis as Eve, all the light will be directed to

one of his detectors, due to the interference. The detector, which is now operating in the linear instead of the Geiger mode (avalanche photon diode), will click and Eve will now share the same key bit with Bob. If Bob chooses to measure in the complementary basis, the light will be divided in two components and its intensity will not be enough to trigger neither of the detectors; therefore, Bob will not get a click and this iteration will be discarded. Subsequently, Alice and Bob will keep for the key the bits for which Alice's preparation basis and Bob's measuring basis agree. During their classical communication, Eve will learn exactly which are these bits; therefore, she will share the same key, while her interference remains unnoticed.

This attack is ineffective for the case of our first, two-way, protocol, in which only Alice is performing the measurement on the pulses received back from Bob. Note that her performing the inverse QW, followed by the measurement in the computational basis, is equivalent to measuring in an *unknown* to Eve (ensured by our Holevo argument mentioned in Sect. 3.1, and presented with details in [106]), "rotated" basis, with respect to the computational one. Therefore, virtually all Eve attempts to perform the detector blinding attack would result in no detection events for Alice. Moreover, even the (rare) detections, being uncorrelated with the initial state sent by Alice, would not pass the verification procedure described in Sect. 3.1.1, as well as the analogous checking rounds of the third, two-way semi-quantum protocol (when Bob reflects the pulses back to Alice and she performs the inverse walk and measures in the computational basis).

Regarding our second, one-way key distribution protocol, Bob's action is similar to the one in the standard protocols: he measures in one of the two publicly known bases. To counter such an attack, Bob can apply one of the known countermeasures proposed and analyzed in [34,59,60,101,108]. Nevertheless, we would like to note again that our QW protocol is more complex than the standard ones based on few (typically four) quantum states, and thus, its implementations might possibly invoke new challenges, a topic worth a separate study.

Time-shift attacks take advantage of the different timing responses of the detectors. Assuming Eve knows the timings of during which each detector is (in)sensitive, allows her to, similarly as in the previous case of the detector blinding, enforce the particular outcomes of Bob's/Alice's measurements. Analogously as in the case of detector blinding attacks, such strategy cannot pass the two-way verification procedures and checking rounds of our first and third protocols. In the case of our one-way protocol, one could employ similar methods to the ones proposed in [57,73,84,118], in order to defend against a time-shift attack.

7 Conclusions and future work

In this paper, we employed, for the first time to our knowledge, QWs in order to design and analyze new secure QKD protocols.

Perhaps the most important contribution of this work is that it introduces the exciting possibility of using QWs for QKD purposes and may spur new research in both cryptography, and also in QWs. By themselves, QWs exhibit many fascinating properties which, as we've shown here, translate to interesting properties of QKD protocols.

Besides the theoretically interesting intersection of two unique and fascinating fields of quantum information science, there are also potential practical benefits in pursuing this investigation. Recently, it has been shown in numerous studies [8,18,20–22,80,81,100] that when Alice and Bob exchange higher dimensional systems (qudits instead of qubits), the respective QKD protocols (in fact both kinds of prepare-and-measure and entanglement-based) *tolerate more noise* than the 2-dimensional ones. This is also confirmed by our results for the high-dimensional states generated by means of QWs, where the noise tolerance increases for higher dimensions. While such systems hold interesting theoretical properties, it could also be that, in a future quantum infrastructure, the generation of these QW states would be easier compared to other higher dimensional systems. Indeed, producing such states may not need the high entanglement of many qubits—instead they could be generated through the evolution of a single-qubit walker on, for instance, a multi-node quantum network.

In what follows, we point out some directions of future work. First, it would be interesting to perform a more detailed study on the two verification procedures presented in Sects. 3.1.1 and 3.1.2 and compare them with respect to various attack strategies. Moreover, one could analyze the relation between the two for concrete cases of Eve's cheating strategies in the presence of noise.

We have presented our protocols in the case that the parties use lossless channels, which is the common first step, when it comes to proving security of new theoretical QKD protocols [7,22,66,87,95]. A relevant future direction of work would be to conduct a detailed analysis of the effects of lossy channels on the number of secret key bits that the parties share in the end, and determine the relevant measures, such as distance versus key-rate, etc.

In Sect. 4, we proved the security of the one-way protocol, but still some improvements could be done. In particular, one could find an analytical solution for the optimal choice of QW parameters or, alternatively, given particular QW parameters, to find an analytical solution for the value of c from Eq. (9). Another interesting question would be to understand the maximally tolerated noise as the dimension of the QW $P \to \infty$ (and also $T_{\max} \to \infty$). For instance, in [22], a high-dimensional QKD protocol was introduced (not using QWs, but simpler states), which could suffer a disturbance up to 50% as the dimension of the state sent by Alice approached infinity (for a novel protocol for multiparty quantum key management that implements d-level measurements, see [114]). Can we construct a QW-QKD protocol with similar features? Does our protocol approach this disturbance level for high P? Moreover, studying and employing other QW models (perhaps the memory-based QWs described and analyzed in, e.g., [1,17,37,50,77,88]) or QWs on different graphs, would be interesting—our key-rate equation would generalize to these cases, the only change would be the value of c; perhaps different QW models, or different graphs, would produce more optimal values, thus increasing the key-rate.

The SQKD protocol we proposed lacks of a proof of security beyond robustness. As we already mentioned in Sect. 5, this proof is technically very challenging due to the high-dimensional QW states and the use of a two-way channel. Hence, computing analytically the key-rate is extremely hard and moreover, the numerical simulation is equally challenging, even for low-dimensional walks. Nevertheless, we believe that obtaining the key rate is not impossible, and we expect that this analysis will yield

quite high error-tolerance. A first step toward this direction would be to try to reduce this protocol to a simpler one (for instance, the one in [14], for which there is a security proof [51]) and prove that it is at least as secure. This reduction does not seem to be a straightforward task and requires a thorough analysis.

Finally, concrete proposals for possible practical implementations of our protocols deserve a separate investigation, and we leave them as future work. While such practical implementations can be non-trivial, we believe that our theoretical proposals are quite promising. In particular, we are optimistic that the experimental advances in the field of the generation and manipulation of high-dimensional states mentioned above, combined with the on-going progress in the field of QW realisations, will permit the implementation of our protocols in higher dimensions. Furthermore, as we argued before, high-dimensional QW states (at least in the case of our second protocol) are not necessary in order to obtain a secure and robust protocol. Along these lines, our proposals could be seen as alternatives to the already existing QKD protocols, which can be implemented with the current technology. Perhaps, the most relevant setups for QKD purposes would be the optical implementations of QWs [16,19,39,40,48,90,93,97,98]. For a detailed and complete review of QW experimental realisations, we refer to the book [75].

Acknowledgements WK would like to acknowledge the hospitality of SQIG–Security and Quantum Information Group in IT—Instituto de Telecomunicações, in Lisbon, during his visit while working on this project. CV acknowledges the support from DP-PMI and FCT (Portugal) through the Grant PD/BD/ 52652/2014. CV, PM, NP and AS acknowledge the support of SQIG-Security and Quantum Information Group. PM, NP and AS also acknowledge the support from UID/EEA/50008/2013 and the support of the project Quantum-Mining POCI-01-0145-FEDER-031826 funded by FCT. NP acknowledges the IT project QbigD funded by FCT PEst-OE/ EEI/LA0008/2013. PM and AS acknowledges the FCT project Confident PTDC/EEI-CTP/4503/2014. A.S. also acknowledges the support of LASIGE Research Unit, ref. UID/CEC/00408/2013.

References

1. Aharonov, D., Ambainis, A., Kempe, J., Vazirani, U.: Quantum walks on graphs. In: Proceedings of ACM Symposium on Theory of Computation (STOC '01), pp. 50–59 (2001)
2. Aharonov, Y., Davidovich, L., Zagury, N.: Quantum random walks. Phys. Rev. A **48**, 1687–1690 (1993). https://doi.org/10.1103/PhysRevA.48.1687
3. Ambainis, A.: Quantum walks and their algorithmic applications. Int. J. Quantum Inf. **01**(04), 507–518 (2003). https://doi.org/10.1142/S0219749903000383
4. Ashwin, N., Ashvin, V.: Quantum walk on the line. Tech. Rep. arXiv:quant-ph/0010117 (2000)
5. Bae, J., Acín, A.: Key distillation from quantum channels using two-way communication protocols. Phys. Rev. A **75**, 012,334 (2007). https://doi.org/10.1103/PhysRevA.75.012334
6. Bavaresco, J., Herrera Valencia, N., Klckl, C., Pivoluska, M., Friis, N., Malik, M., Huber, M.: Two measurements are sufficient for certifying high-dimensional entanglement. arXiv:1709.07344 (2017)
7. Beaudry, N.J., Lucamarini, M., Mancini, S., Renner, R.: Security of two-way quantum key distribution. Phys. Rev. A **88**(6), 062,302 (2013)
8. Bechmann-Pasquinucci, H., Tittel, W.: Quantum cryptography using larger alphabets. Phys. Rev. A **61**, 062,308 (2000). https://doi.org/10.1103/PhysRevA.61.062308
9. Bedington, R., Arrazola, J.M., Ling, A.: Progress in satellite quantum key distribution. Npj Quantum Inf. **3**, 30 (2017). https://doi.org/10.1038/s41534-017-0031-5
10. Bennett, C., Brassard, G.: Quantum cryptography: public key distribution and coin tossing. In: Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, pp. 175–179. IEEE Press, New York (1984)

11. Bennett, C.H., Brassard, G., Mermin, N.D.: Quantum cryptography without bell's theorem. Phys. Rev. Lett. **68**, 557–559 (1992). https://doi.org/10.1103/PhysRevLett.68.557
12. Berta, M., Christandl, M., Colbeck, R., Renes, J.M., Renner, R.: The uncertainty principle in the presence of quantum memory. Nat. Phys. **6**, 659–662 (2010). https://doi.org/10.1038/nphys1734
13. Boyer, M., Gelles, R., Kenigsberg, D., Mor, T.: Semiquantum key distribution. Phys. Rev. A **79**, 032,341 (2009). https://doi.org/10.1103/PhysRevA.79.032341
14. Boyer, M., Kenigsberg, D., Mor, T.: Quantum key distribution with classical bob. Phys. Rev. Lett. **99**, 140,501 (2007). https://doi.org/10.1103/PhysRevLett.99.140501
15. Brassard, G., Lütkenhaus, N., Mor, T., Sanders, B.: Security aspects of practical quantum cryptography. In: In Advances in Cryptology? EUROCRYPT'2000, pp. 289–299 (2000)
16. Broome, M.A., Fedrizzi, A., Lanyon, B.P., Kassal, I., Aspuru-Guzik, A., White, A.G.: Discrete single-photon quantum walks with tunable decoherence. Phys. Rev. Lett. **104**, 153,602 (2010). https://doi.org/10.1103/PhysRevLett.104.153602
17. Brun, T.A., Carteret, H.A., Ambainis, A.: Quantum walks driven by many coins. Phys. Rev. A **67**, 052,317 (2003). https://doi.org/10.1103/PhysRevA.67.052317
18. Bruss, D., Christandl, M., Ekert, A., Englert, B.G., Kaszlikowski, D., Macchiavello, C.: Tomographic quantum cryptography: equivalence of quantum and classical key distillation. Phys. Rev. Lett. **91**, 097,901 (2003). https://doi.org/10.1103/PhysRevLett.91.097901
19. Cardano, F., Massa, F., Qassim, H., Karimi, E., Slussarenko, S., Paparo, D., de Lisio, C., Sciarrino, F., Santamato, E., Boyd, R.W., Marrucci, L.: Quantum walks and wavepacket dynamics on a lattice with twisted photons. Sci. Adv. **1**, e1500,087 (2015). https://doi.org/10.1126/sciadv.1500087
20. Cerf, N.J., Bourennane, M., Karlsson, A., Gisin, N.: Security of quantum key distribution using d-level systems. Phys. Rev. Lett. **88**, 127,902 (2002). https://doi.org/10.1103/PhysRevLett.88.127902
21. Chau, H.F.: Unconditionally secure key distribution in higher dimensions by depolarization. IEEE Trans. Inf. Theory **51**, 1451–1468 (2005). https://doi.org/10.1109/TIT.2005.844076
22. Chau, H.F.: Quantum key distribution using qudits that each encode one bit of raw key. Phys. Rev. A **92**, 062,324 (2015). https://doi.org/10.1103/PhysRevA.92.062324
23. Childs, A., Cleve, R., Deotto, E., Farhi, E., Gutmann, S., Spielman, D.: Exponential algorithmic speedup by a quantum walk. In: Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, STOC '03, pp. 59–68. ACM, New York, NY, USA (2003)
24. Childs, A.M.: Universal computation by quantum walk. Phys. Rev. Lett. **102**, 180,501 (2009). https://doi.org/10.1103/PhysRevLett.102.180501
25. Childs, A.M., Gosset, D., Webb, Z.: Universal computation by multiparticle quantum walk. Science **339**(6121), 791–794 (2013). https://doi.org/10.1126/science.1229957
26. Christandl, M., König, R., Renner, R.: Postselection technique for quantum channels with applications to quantum cryptography. Phys. Rev. Lett. **102**, 020,504 (2009). https://doi.org/10.1103/PhysRevLett.102.020504
27. Curty, M., Guhne, O., Lewenstein, M., Lutkenhaus, N.: Detecting two-party quantum correlations in quantum-key-distribution protocols. Phys. Rev. A **71**, 022,306 (2005). https://doi.org/10.1103/PhysRevA.71.022306
28. Curty, M., Lewenstein, M., Lutkenhaus, N.: Entanglement as a precondition for secure quantum key distribution. Phys. Rev. Lett. **92**, 217,903 (2004). https://doi.org/10.1103/PhysRevLett.92.217903
29. Dada, A.C., Leach, J., Buller, G.S., Padgett, M.J., Andersson, E.: Experimental high-dimensional two-photon entanglement and violations of generalized bell inequalities. Nat. Phys. **7**, 677–680 (2011). https://doi.org/10.1038/nphys1996
30. Devetak, I., Winter, A.: Distillation of secret key and entanglement from quantum states. Proc. R. Soc. Lond. A Math. Phys. Eng. Sci. **461**(2053), 207–235 (2005). https://doi.org/10.1098/rspa.2004.1372
31. Diamanti, E., Lo, H.K., Qi, B., Yuan, Z.: Practical challenges in quantum key distribution. Npj Quantum Inf. **2**, 16,025 (2016). https://doi.org/10.1038/npjqi.2016.25
32. Dixon, A.R., Dynes, J.F., Lucamarini, M., Frhlich, B., Sharpe, A.W., Plews, A., Tam, W., Yuan, Z.L., Tanizawa, Y., Sato, H., Kawamura, S., Fujiwara, M., Sasaki, M., Shields, A.J.: Quantum key distribution with hacking countermeasures and long term field trial. Sci. Rep. **7**, 1978 (2017). https://doi.org/10.1038/s41598-017-01884-0
33. Ekert, A.K.: Quantum cryptography based on bell's theorem. Phys. Rev. Lett. **67**, 661–663 (1991). https://doi.org/10.1103/PhysRevLett.67.661

34. Elezov, M., Ozhegov, R., Kurochkin, Y., Goltsman, G., Makarov, V.: Countermeasures against blinding attack on superconducting nanowire detectors for QKD. EPJ Web Conf. **103**, 10,002 (2015). https://doi.org/10.1051/epjconf/201510310002
35. Erhard, M., Malik, M., Krenn, M., Zeilinger, A.: Experimental ghz entanglement beyond qubits. arXiv:1708.03881 (2017)
36. Farhi, E., Gutmann, S.: Quantum computation and decision trees. Phys. Rev. A **58**, 915–928 (1998). https://doi.org/10.1103/PhysRevA.58.915
37. Gettrick, M., Miszczak, J.A.: Quantum walks with memory on cycles. Phys. A Stat. Mech. Its Appl. **399**, 163–170 (2014). https://doi.org/10.1016/j.physa.2014.01.002
38. Gisin, N., Fasel, S., Kraus, B., Zbinden, H., Ribordy, G.: Trojan-horse attacks on quantum-key-distribution systems. Phys. Rev. A **73**, 022,320 (2006). https://doi.org/10.1103/PhysRevA.73.022320
39. Goyal, S.K., Roux, F.S., Forbes, A., Konrad, T.: Implementing quantum walks using orbital angular momentum of classical light. Phys. Rev. Lett. **110**, 263,602 (2013). https://doi.org/10.1103/PhysRevLett.110.263602
40. Goyal, S.K., Roux, F.S., Forbes, A., Konrad, T.: Implementation of multidimensional quantum walks using linear optics and classical light. Phys. Rev. A **92**, 040,302 (2015). https://doi.org/10.1103/PhysRevA.92.040302
41. Harrington, J.W., Ettinger, J.M., Hughes, R.J., Nordholt, J.E.: Enhancing practical security of quantum key distribution with a few decoy states. arXiv:quant-ph/0503002 (2005)
42. Hiesmayr, B.C., de Dood, M.J.A., Löffler, W.: Observation of four-photon orbital angular momentum entanglement. Phys. Rev. Lett. **116**, 073,601 (2016). https://doi.org/10.1103/PhysRevLett.116.073601
43. Hwang, W.Y.: Quantum key distribution with high loss: toward global secure communication. Phys. Rev. Lett. **91**, 057,901 (2003). https://doi.org/10.1103/PhysRevLett.91.057901
44. Islam, N.T., Cahall, C., Aragoneses, A., Lezama, A., Kim, J., Gauthier, D.J.: Robust and stable delay interferometers with application to d-dimensional time-frequency quantum key distribution. Phys. Rev. Appl. **7**, 044,010 (2017). https://doi.org/10.1103/PhysRevApplied.7.044010
45. Jain, N., Stiller, B., Khan, I., Elser, D., Marquardt, C., Leuchs, G.: Attacks on practical quantum key distribution systems (and how to prevent them). Contemp. Phys. **57**(3), 366–387 (2016). https://doi.org/10.1080/00107514.2016.1148333
46. Jha, A.K., Malik, M., Boyd, R.W.: Exploring energy-time entanglement using geometric phase. Phys. Rev. Lett. **101**, 180,405 (2008). https://doi.org/10.1103/PhysRevLett.101.180405
47. Kempe, J.: Quantum random walks: an introductory overview. Contemp. Phys. **44**(4), 307–327 (2003). https://doi.org/10.1080/0010715103100011077 6
48. Knight, P.L., Roldan, E., Sipe, J.E.: Quantum walk on the line as an interference phenomenon. Phys. Rev. A **68**, 020,301 (2003). https://doi.org/10.1103/PhysRevA.68.020301
49. Krawec, W.O.: Restricted attacks on semi-quantum key distribution protocols. Quantum Inf. Process. **13**(11), 2417–2436 (2014). https://doi.org/10.1007/s11128-014-0802-2
50. Krawec, W.O.: History dependent quantum walk on the cycle with an unbalanced coin. Phys. A Stat. Mech. Its Appl. **428**, 319–331 (2015). https://doi.org/10.1016/j.physa.2015.02.061
51. Krawec, W.O.: Security proof of a semi-quantum key distribution protocol. In: 2015 IEEE International Symposium on Information Theory (ISIT), pp. 686–690. IEEE (2015)
52. Krawec, W.O.: Quantum key distribution with mismatched measurements over arbitrary channels. arXiv:1608.07728 (2016)
53. Krawec, W.O.: Security of a semi-quantum protocol where reflections contribute to the secret key. Quantum Inf. Process. **15**(5), 2067–2090 (2016)
54. Krenn, M., Huber, M., Fickler, R., Lapkiewicz, R., Ramelow, S., Zeilinger, A.: Generation and confirmation of a (100×100)-dimensional entangled quantum system. Proc. Natl. Acad. Sci. USA **111**, 6243–6247 (2014). https://doi.org/10.1073/pnas.1402365111
55. Krenn, M., Malik, M., Fickler, R., Lapkiewicz, R., Zeilinger, A.: Automated search for new quantum experiments. Phys. Rev. Lett. **116**, 090,405 (2016). https://doi.org/10.1103/PhysRevLett.116.090405
56. Kurtsiefer, C., Zarda, P., Mayer, S., Weinfurter, H.: The breakdown flash of silicon avalanche photodiodes-back door for eavesdropper attacks? J. Mod. Opt. **48**(13), 2039–2047 (2001). https://doi.org/10.1080/09500340108240905
57. Lamas-Linares, A., Kurtsiefer, C.: Breaking a quantum key distribution system through a timing side channel. Opt. Express **15**, 9388–9393 (2007). https://doi.org/10.1364/OE.15.009388

58. Lavery, M.P.J., Robertson, D.J., Berkhout, G.C.G., Love, G.D., Padgett, M.J., Courtial, J.: Refractive elements for the measurement of the orbital angular momentum of a single photon. Opt. Express **20**, 2110–2115 (2012). https://doi.org/10.1364/OE.20.002110
59. Lee, M.S., Park, B.K., Woo, M.K., Park, C.H., Kim, Y.S., Han, S.W., Moon, S.: Countermeasure against blinding attacks on low-noise detectors with a background-noise-cancellation scheme. Phys. Rev. A **94**, 062,321 (2016). https://doi.org/10.1103/PhysRevA.94.062321
60. Liu, Q., Lamas-Linares, A., Kurtsiefer, C., Skaar, J., Makarov, V., Gerhardt, I.: A universal setup for active control of a single-photon detector. Rev. Sci. Instrum. **85**(1), 013,108 (2014). https://doi.org/10.1063/1.4854615
61. Lizama-Prez, L.A., Lpez, J.M., De Carlos Lpez, E.: Quantum key distribution in the presence of the intercept-resend with faked states attack. Entropy **19**(1), 4 (2016). https://doi.org/10.3390/e19010004
62. Lo, H.K., Chau, H.: Unconditional security of quantum key distribution over arbitrarily long distances. Science **283**(5410), 2050–2056 (1999). https://doi.org/10.1126/science.283.5410.2050. Please check and confirm the inserted publication year is correct for the references [61, 68]
63. Lo, H.K., Chau, H.F., Ardehali, M.: Efficient quantum key distribution scheme and a proof of its unconditional security. J. Cryptol. **18**(2), 133–165 (2005). https://doi.org/10.1007/s00145-004-0142-y
64. Lo, H.K., Ma, X., Chen, K.: Decoy state quantum key distribution. Phys. Rev. Lett. **94**, 230,504 (2005). https://doi.org/10.1103/PhysRevLett.94.230504
65. Lovett, N.B., Cooper, S., Everitt, M., Trevers, M., Kendon, V.: Universal quantum computation using the discrete-time quantum walk. Phys. Rev. A **81**, 042,330 (2010). https://doi.org/10.1103/PhysRevA.81.042330
66. Lu, H., Fung, C.H.F., Ma, X., Cai, Q.Y.: Unconditional security proof of a deterministic quantum key distribution with a two-way quantum channel. Phys. Rev. A **84**, 042,344 (2011). https://doi.org/10.1103/PhysRevA.84.042344
67. Lucamarini, M., Choi, I., Ward, M.B., Dynes, J.F., Yuan, Z.L., Shields, A.J.: Practical security bounds against the trojan-horse attack in quantum key distribution. Phys. Rev. X **5**, 031,030 (2015). https://doi.org/10.1103/PhysRevX.5.031030
68. Lucamarini, M., Dynes, J.F., Frohlich, B., Yuan, Z., Shields, A.J.: Security bounds for efficient decoy-state quantum key distribution. IEEE J. Sel. Top. Quantum Electron. **21**(3), 1–8 (2015). https://doi.org/10.1109/JSTQE.2015.2394774
69. Lutkenhaus, N.: Security against individual attacks for realistic quantum key distribution. Phys. Rev. A **61**, 052,304 (2000). https://doi.org/10.1103/PhysRevA.61.052304
70. Lutkenhaus, N., Jahma, M.: Quantum key distribution with realistic states: photon-number statistics in the photon-number splitting attack. New J. Phys. **4**(1), 44 (2002)
71. Lydersen, L., Wiechers, C., Wittmann, C., Elser, D., Skaar, J., Makarov, V.: Hacking commercial quantum cryptography systems by tailored bright illumination. Nat. Photon. **4**, 686 (2010). https://doi.org/10.1038/nphoton.2010.214
72. Ma, X., Qi, B., Zhao, Y., Lo, H.K.: Practical decoy state for quantum key distribution. Phys. Rev. A **72**, 012,326 (2005). https://doi.org/10.1103/PhysRevA.72.012326
73. Makarov, V., Anisimov, A., Skaar, J.: Effects of detector efficiency mismatch on security of quantum cryptosystems. Phys. Rev. A **74**, 022,313 (2006). https://doi.org/10.1103/PhysRevA.74.022313
74. Malik, M., Erhard, M., Huber, M., Krenn, M., Fickler, R., Zeilinger, A.: Multi-photon entanglement in high dimensions. Nat. Photon. **10**, 248–252 (2016). https://doi.org/10.1038/nphoton.2016.12
75. Manouchehri, K., Wang, J.: Physical Implementation of Quantum Walks. Springer Publishing Company, Incorporated, New York (2013)
76. Martin, A., Guerreiro, T., Tiranov, A., Designolle, S., Frowis, F., Brunner, N., Huber, M., Gisin, N.: Quantifying photonic high-dimensional entanglement. Phys. Rev. Lett. **118**, 110,501 (2017). https://doi.org/10.1103/PhysRevLett.118.110501
77. McGettrick, M.: One dimensional quantum walks with memory. Quantum Inf. Comput. **10**(5), 509–524 (2010)
78. Mirhosseini, M., Malik, M., Shi, Z., Boyd, R.W.: Efficient separation of the orbital angular momentum eigenstates of light. Nat. Commun. **4**, 2781 (2013). https://doi.org/10.1038/ncomms3781
79. Nikolopoulos, G.: Applications of single-qubit rotations in quantum public-key cryptography. Phys. Rev. A **77**, 032,348 (2008). https://doi.org/10.1103/PhysRevA.77.032348
80. Nikolopoulos, G.M., Alber, G.: Security bound of two-basis quantum-key-distribution protocols using qudits. Phys. Rev. A **72**, 032,320 (2005). https://doi.org/10.1103/PhysRevA.72.032320

81. Nikolopoulos, G.M., Ranade, K.S., Alber, G.: Error tolerance of two-basis quantum-key-distribution protocols using qudits and two-way classical communication. Phys. Rev. A **73**, 032,325 (2006). https://doi.org/10.1103/PhysRevA.73.032325
82. Portugal, R.: Quantum Walks and Search Algorithms. Quantum Science and Technology. Springer, New York, NY (2013). https://doi.org/10.1007/978-1-4614-6336-8
83. Pugh, C.J., Kaiser, S., Bourgoin, J.P., Jin, J., Sultana, N., Agne, S., Anisimova, E., Makarov, V., Choi, E., Higgins, B.L., Jennewein, T.: Airborne demonstration of a quantum key distribution receiver payload. Quantum Sci. Technol. **2**, 024,009 (2017). https://doi.org/10.1088/2058-9565/aa701f
84. Qi, B., Fung, C.H.F., Lo, H.K., Ma, X.: Time-shift attack in practical quantum cryptosystems. Quantum Inf. Comput. **7**, 73–82 (2007). https://doi.org/10.1103/PhysRevX.5.031030
85. Quantique, I.: Quantum key distribution record (2017). https://www.idquantique.com/clavis3-quantum-key-distribution-platform/. Accessed 12 Sept 2018
86. Renner, R.: Symmetry of large physical systems implies independence of subsystems. Nat. Phys. **3**, 645–649 (2007). https://doi.org/10.1038/nphys684
87. Renner, R., Gisin, N., Kraus, B.: Information-theoretic security proof for quantum-key-distribution protocols. Phys. Rev. A **72**, 012,332 (2005). https://doi.org/10.1103/PhysRevA.72.012332
88. Rohde, P.P., Brennen, G.K., Gilchrist, A.: Quantum walks with memory provided by recycled coins and a memory of the coin-flip history. Phys. Rev. A **87**, 052,302 (2013). https://doi.org/10.1103/PhysRevA.87.052302
89. Rohde, P.P., Fitzsimons, J.F., Gilchrist, A.: Quantum walks with encrypted data. Phys. Rev. Lett. **109**(15), 150,501 (2012)
90. Roldan, E., Soriano, J.C.: Optical implementability of the two-dimensional quantum walk. J. Mod. Opt. **52**, 2649–2657 (2006). https://doi.org/10.1080/09500340500309873
91. Rosenberg, D., Peterson, C.G., Harrington, J.W., Rice, P.R., Dallmann, N., Tyagi, K.T., McCabe, K.P., Nam, S., Baek, B., Hadfield, R.H., Hughes, R.J., Nordholt, J.E.: Practical long-distance quantum key distribution system using decoy levels. New J. Phys. **11**(4), 045,009 (2009)
92. Sajeed, S., Minshull, C., Jain, N., Makarov, V.: Invisible Trojan-horse attack. Sci. Rep. **7**, 8403 (2017). https://doi.org/10.1038/s41598-017-08279-1
93. Sansoni, L., Sciarrino, F., Vallone, G., Mataloni, P., Crespi, A., Ramponi, R., Osellame, R.: Two-particle bosonic-fermionic quantum walk via integrated photonics. Phys. Rev. Lett. **108**, 010,502 (2012). https://doi.org/10.1103/PhysRevLett.108.010502
94. Santha, M.: Quantum walk based search algorithms. In: Agrawal, M., Du, D., Duan, Z., Li, A. (eds.) Theory and Applications of Models of Computation, Lecture Notes in Computer Science, vol. 4978, pp. 31–46. Springer, Berlin (2008). https://doi.org/10.1007/978-3-540-79228-4-3
95. Scarani, V., Bechmann-Pasquinucci, H., Cerf, N., Dušek, M., Lutkenhaus, N., Peev, M.: The security of practical quantum key distribution. Rev. Mod. Phys. **81**, 1301–1350 (2009). https://doi.org/10.1103/RevModPhys.81.1301
96. Scarani, V., Bechmann-Pasquinucci, H., Cerf, N.J., Dušek, M., Lütkenhaus, N., Peev, M.: The security of practical quantum key distribution. Rev. Mod. Phys. **81**, 1301–1350 (2009). https://doi.org/10.1103/RevModPhys.81.1301
97. Schreiber, A., Cassemiro, K.N., Potocek, V., Gabris, A., Mosley, P.J., Andersson, E., Jex, I., Silberhorn, C.: Photons walking the line: a quantum walk with adjustable coin operations. Phys. Rev. Lett. **104**, 050,502 (2010). https://doi.org/10.1103/PhysRevLett.104.050502
98. Schreiber, A., Gabris, A., Rohde, P.P., Laiho, K., Stefanak, M., Potocek, V., Hamilton, C., Jex, I., Silberhorn, C.: A 2D quantum walk simulation of two-particle dynamics. Science **336**, 55–58 (2012). https://doi.org/10.1126/science.1218448
99. Seyfarth, U., Nikolopoulos, G., Alber, G.: Symmetries and security of a quantum-public-key encryption based on single-qubit rotations. Phys. Rev. A **85**(2), 022,342 (2012)
100. Sheridan, L., Scarani, V.: Security proof for quantum key distribution using qudit systems. Phys. Rev. A **82**, 030,301 (2010). https://doi.org/10.1103/PhysRevA.82.030301
101. Stipcević, M.: Preventing detector blinding attack and other random number generator attacks on quantum cryptography by use of an explicit random number generator. arXiv:1403.0143 (2014)
102. Takesue, H., Sasaki, T., Tamaki, K., Koashi, M.: Experimental quantum key distribution without monitoring signal disturbance. Nat. Photon. **9**, 827–831 (2015). https://doi.org/10.1038/nphoton.2015.173
103. Ursin, R., Tiefenbacher, F., Schmitt-Manderbach, T., Weier, H., Scheidl, T., Lindenthal, M., Blauensteiner, B., Jennewein, T., Perdigues, J., Trojek, P., Omer, B., Furst, M., Meyenburg, M., Rarity, J.,

Sodnik, Z., Barbieri, C., Weinfurter, H., Zeilinger, A.: Entanglement-based quantum communication over 144 km. Nat. Phys. **3**, 481–486 (2007). https://doi.org/10.1038/nphys629
104. Vakhitov, A., Makarov, V., Hjelme, D.R.: Large pulse attack as a method of conventional optical eavesdropping in quantum cryptography. J. Mod. Opt. **48**(13), 2023–2038 (2001). https://doi.org/10.1080/09500340108240904
105. Venegas-Andraca, S.E.: Quantum walks: a comprehensive review. Quantum Inf. Process. **11**(5), 1015–1106 (2012). https://doi.org/10.1007/s11128-012-0432-5
106. Vlachou, C., Rodrigues, J., Mateus, P., Paunković, N., Souto, A.: Quantum walk public-key cryptographic system. Int. J. Quantum Inf. **13**(07), 1550,050 (2015)
107. Wang, F., Erhard, M., Babazadeh, A., Malik, M., Krenn, M., Zeilinger, A.: Generation of the complete four-dimensional bell basis. Optica **4**, 1462–1467 (2017). https://doi.org/10.1364/OPTICA.4.001462
108. Wang, J., Wang, H., Qin, X., Wei, Z., Zhang, Z.: The countermeasures against the blinding attack in quantum key distribution. Eur. Phys. J. D **70**(1), 5 (2016). https://doi.org/10.1140/epjd/e2015-60469-8
109. Wang, Q., Wang, X.B., Bjrk, G., Karlsson, A.: Improved practical decoy state method in quantum key distribution with parametric down-conversion source. EPL Europhys. Lett. **79**(4), 40,001 (2007)
110. Wang, X.B.: Beating the photon-number-splitting attack in practical quantum cryptography. Phys. Rev. Lett. **94**, 230,503 (2005). https://doi.org/10.1103/PhysRevLett.94.230503
111. Wang, X.B.: Decoy-state protocol for quantum cryptography with four different intensities of coherent light. Phys. Rev. A **72**, 012,322 (2005). https://doi.org/10.1103/PhysRevA.72.012322
112. Wiechers, C., Lydersen, L., Wittmann, C., Elser, D., Skaar, J., Marquardt, C., Makarov, V., Leuchs, G.: After-gate attack on a quantum cryptosystem. New J. Phys. **13**(1), 013043 (2011)
113. Woodhead, E., Pironio, S.: Secrecy in prepare-and-measure clauser-horne-shimony-holt tests with a qubit bound. Phys. Rev. Lett. **115**, 150,501 (2015). https://doi.org/10.1103/PhysRevLett.115.150501
114. Xu, G., Chen, X.B., Dou, Z., Yang, Y.X., Li, Z.: A novel protocol for multiparty quantum key management. Quantum Inf. Process. **14**(8), 2959–2980 (2015). https://doi.org/10.1007/s11128-015-1021-1
115. Yao, A.C.: How to generate and exchange secrets. In: 27th Annual Symposium on Foundations of Computer Science, 1986., pp. 162–167. IEEE (1986)
116. Yin, J., Cao, Y., Li, Y.H., Liao, S.K., Zhang, L., Ren, J.G., Cai, W.Q., Liu, W.Y., Li, B., Dai, H., Li, G.B., Lu, Q.M., Gong, Y.H., Xu, Y., Li, S.L., Li, F.Z., Yin, Y.Y., Jiang, Z.Q., Li, M., Jia, J.J., Ren, G., He, D., Zhou, Y.L., Zhang, X.X., Wang, N., Chang, X., Zhu, Z.C., Liu, N.L., Chen, Y.A., Lu, C.Y., Shu, R., Peng, C.Z., Wang, J.Y., Pan, J.W.: Satellite-based entanglement distribution over 1200 kilometers. Science **356**, 1140–1144 (2017). https://doi.org/10.1126/science.aan3211
117. Zhang, W., Qiu, D., Mateus, P.: Security of a single-state semi-quantum key distribution protocol. arXiv:1612.03170 (2016)
118. Zhao, Y., Fung, C.H.F., Qi, B., Chen, C., Lo, H.K.: Quantum hacking: experimental demonstration of time-shift attack against practical quantum-key-distribution systems. Phys. Rev. A **78**, 042,333 (2008). https://doi.org/10.1103/PhysRevA.78.042333
119. Zou, X., Qiu, D., Li, L., Wu, L., Li, L.: Semiquantum-key distribution using less than four quantum states. Phys. Rev. A **79**, 052,312 (2009). https://doi.org/10.1103/PhysRevA.79.052312

Affiliations

Chrysoula Vlachou[1,2] · Walter Krawec[4] · Paulo Mateus[1,2] · Nikola Paunković[1,2,3] · André Souto[1,5,6]

Chrysoula Vlachou
chrysoula.vlachou@tecnico.ulisboa.pt

Walter Krawec
walter.krawec@uconn.edu

Paulo Mateus
pmat@math.ist.utl.pt

Nikola Paunković
npaunkov@math.ist.utl.pt

1 Instituto de Telecomunicações, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal

2 Dep. de Matemática – Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal

3 CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal

4 Computer Science and Engineering Department, University of Connecticut, Storrs, CT 06268, USA

5 Dep. de Informática – Faculdade de Ciências, Universidade de Lisboa, Campo Grande 016, 1749-016 Lisbon, Portugal

6 LASIGE – Faculdade de Ciências, Universidade de Lisboa, Campo Grande 016, 1749-016 Lisbon, Portugal

Uhlmann Connection in Fermionic Systems Undergoing Phase Transitions

Bruno Mera,[1,2,3] Chrysoula Vlachou,[4,3] Nikola Paunković,[4,3] and Vítor R. Vieira[1,2]
[1]CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
[2]Departamento de Física, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
[3]Instituto de Telecomunicações, 1049-001 Lisbon, Portugal
[4]Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
(Received 23 February 2017; revised manuscript received 15 May 2017; published 7 July 2017)

We study the behavior of the Uhlmann connection in systems of fermions undergoing phase transitions. In particular, we analyze some of the paradigmatic cases of topological insulators and superconductors in one dimension, as well as the BCS theory of superconductivity in three dimensions. We show that the Uhlmann connection signals phase transitions in which the eigenbasis of the state of the system changes. Moreover, using the established fidelity approach and the study of the edge states, we show the absence of thermally driven phase transitions in the case of topological insulators and superconductors. We clarify what is the relevant parameter space associated with the Uhlmann connection so that it signals the existence of order in mixed states. In addition, the study of Majorana modes at finite temperature opens the possibility of applications in realistic stable quantum memories. Finally, the analysis of the different behavior of the BCS model and the Kitaev chain, with respect to the Uhlmann connection, suggested that in realistic scenarios the gap of topological superconductors could also, generically, be temperature dependent.

DOI: 10.1103/PhysRevLett.119.015702

Introduction.—Topological phases of matter have been a subject of active research during the last decades, as they constitute a whole new paradigm in condensed matter physics. In contrast to the well-studied standard quantum phases of matter, described by local order parameters (see, for example, Anderson's classification [1]), the ground states of topological systems are globally characterized by topological invariants [2–5]. Hamiltonians of gapped systems with different topological orders cannot be smoothly transformed from one into the other unless passing through a gap-vanishing region of criticality. In particular, insulators and superconductors with an energy gap exhibit topological orders and are classified according to the symmetries that their Hamiltonians possess [6,7], namely, time reversal, particle hole, and chiral symmetry. As opposed to the standard Landau symmetry breaking theory of quantum phase transitions (PTs), in topological PTs the symmetries of the Hamiltonian are not violated. For a topological PT to occur, that is a gapped state of the system to be deformed in another gapped state in a different topological class, the energy gap has to close. In other words, the quantum state of the system undergoing a topological PT is gapless. A manifestation of the topological order of a system is the presence of robust symmetry-protected edge states on the boundary between two distinct topological phases, as predicted by the bulk-to-boundary principle [8].

A question that naturally arises is whether there is any kind of topological order at finite temperatures, and different approaches have been used to tackle this problem [9,10]. One of the most promising approaches is based on the work of Uhlmann [11], who extended the notion of geometrical phases from pure states to density matrices. The concept of the Uhlmann holonomy, and the quantities that can be derived from it, were used to infer PTs at finite temperatures [12–17]. Nevertheless, the physical meaning of these quantities and their relevance to the observable properties of the corresponding systems stay as an interesting open question [18–20]. There exist several proposals for the observation of the Uhlmann geometric phase [21–23] and an experimental realization has been reported in Ref. [24].

Information-theoretic quantities such as entanglement measures [25–27] and the fidelity, a measure of distinguishability between two quantum states [28–32], were extensively used in the study of PTs. Whenever there is a PT, the density matrix of a system changes significantly and, therefore, a sudden drop of the fidelity $F(\rho,\sigma) \equiv \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$ signals out this change. The fidelity, is closely related to the Uhlmann connection, through the Bures metric [33]. Therefore, they can both be used to infer the possibility of PTs, as in Ref. [16] (for the pure-state case of the Berry phase, see Refs. [34,35]).

We analyze the behavior of the fidelity and the Uhlmann connection associated to thermal states in fermionic systems. We consider the space consisting of the parameters of the Hamiltonian and the temperature, as it provides a physically sensible base space for the principal bundle, describing the amplitudes of the density operator. We study paradigmatic models of 1D topological insulators (TIs) (Creutz Ladder [36,37] and Su-Schrieffer-Heeger [38] models) and superconductors (TSCs) (Kitaev chain [39]) with chiral symmetry. We conclude that the effective temperature only smears out the topological features exhibited at zero temperature, without causing a thermal PT. We also analyze the BCS model of superconductivity [40], previously studied in Ref. [16], by

 © 2017 American Physical Society

further identifying the significance of thermal and purely quantum contributions to PTs, using the fidelity and the Uhlmann connection. In contrast to the aforementioned nontrivial topological systems, both quantities indicate the existence of thermal PTs.

This Letter is organized as follows: first, we elaborate on the relationship between the fidelity and the Uhlmann connection and motivate their use in inferring PTs, both at zero and finite temperatures. In the following section, we present our results on the fidelity and the Uhlmann connection for the aforementioned systems and discuss the possibility of temperature driven PTs. In the last section we summarize our conclusions and point out possible directions of future work.

Fidelity and the Uhlmann parallel transport.—Given a Hilbert space, one can consider the set of density matrices with full rank (e.g., thermal states) and the associated set of amplitudes (generalization to the case of the sets of singular density matrices with fixed rank is straightforward). For a state ρ, an associated amplitude w satisfies $\rho = ww^\dagger$. Thus, there exists a unitary (gauge) freedom in the choice of the amplitude, since both w and $w' = wU$, with U being an arbitrary unitary, are associated to the same ρ. Two amplitudes w_1 and w_2, corresponding to states ρ_1 and ρ_2, respectively, are said to be *parallel* in the Uhlmann sense if and only if they minimize the Hilbert-Schmidt distance $\|w_1 - w_2\| = \sqrt{\mathrm{Tr}[(w_1 - w_2)^\dagger (w_1 - w_2)]}$, induced by the inner product $\langle w_1, w_2\rangle = \mathrm{Tr}(w_1^\dagger w_2)$. The condition of parallelism turns out to be equivalent to maximizing $\mathrm{Re}\langle w_1, w_2\rangle$, since $\|w_1 - w_2\|^2 = 2(1 - \mathrm{Re}\langle w_1, w_2\rangle)$. By writing $w_i = \sqrt{\rho_i} U_i$, $i = 1, 2$, where the U_i's are unitary matrices, we get

$$\begin{aligned}\mathrm{Re}\langle w_1, w_2\rangle \le |\langle w_1, w_2\rangle| &= |\mathrm{Tr}(w_2^\dagger w_1)| \\ &= |\mathrm{Tr}(U_2^\dagger \sqrt{\rho_2}\sqrt{\rho_1} U_1)| \\ &= |\mathrm{Tr}(|\sqrt{\rho_2}\sqrt{\rho_1}| U U_1 U_2^\dagger)| \\ &\le \mathrm{Tr}(|\sqrt{\rho_2}\sqrt{\rho_1}|) \\ &= \mathrm{Tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}} = F(\rho_1, \rho_2), \quad (1)\end{aligned}$$

where U is the unitary associated to the polar decomposition of $\sqrt{\rho_2}\sqrt{\rho_1}$, and the penultimate step is the Cauchy-Schwartz inequality. Hence, the equality holds if and only if $U(U_1 U_2^\dagger) = I$. Note that in this case, also the first equality holds, and we have $\mathrm{Re}\langle w_1, w_2\rangle = \langle w_1, w_2\rangle \in \mathbb{R}^+$, which provides yet another interpretation of the Uhlmann parallel transport condition as a generalization of the Berry pure-state connection: the phase, given by $\Phi_U = \arg\langle w_1, w_2\rangle$, is trivial, i.e., zero.

Given a curve of density matrices $\gamma : [0, 1] \ni t \mapsto \rho(t)$ and the initial amplitude $w(0)$ of $\rho(0)$, the Uhlmann parallel transport gives a unique curve of amplitudes $w(t)$ with the property that $w(t)$ is parallel to $w(t + \delta t)$ for an infinitesimal δt (the horizontal lift of γ). The length of this curve of amplitudes, according to the metric induced by the Frobenius inner product, is equal to the length, according to the Bures metric (which is the infinitesimal version of the Bures distance $D_B(\rho_1, \rho_2)^2 = 2[1 - F(\rho_1, \rho_2)]$), of the corresponding curve γ of the density matrices. This shows the relation between the Uhlmann connection and the fidelity (for details, see, for example, Refs. [11,41]).

We see that the "Uhlmann factor" U, given by the polar decomposition $\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)} = |\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)}| U$, characterizes the Uhlmann parallel transport. For two close points t and $t + \delta t$, if the two states $\rho(t)$ and $\rho(t + \delta t)$ belong to the same phase, one expects them to almost commute, resulting in the Uhlmann factor being approximately equal to the identity $\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)} \approx |\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)}|$. On the other hand, if the two states belong to two different phases, one expects them to be drastically different (confirmed by the fidelity approach), both in terms of their eigenvalues and/or eigenvectors, potentially leading to nontrivial $U \neq I$ (see the previous study on the Uhlmann factor and the finite-temperature PTs for the case of the BCS superconductivity [16]). To quantify the difference between the Uhlmann factor and the identity, and thus the nontriviality of the Uhlmann connection, we consider the following quantity:

$$\begin{aligned}\Delta(\rho(t), \rho(t + \delta t)) &:= F(\rho(t), \rho(t + \delta t)) \\ &\quad - \mathrm{Tr}\left[\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)}\right]. \quad (2)\end{aligned}$$

Note that $\Delta = \mathrm{Tr}[|\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)}|(I - U)]$. When the two states are from the same phase we have $\rho(t) \approx \rho(t + \delta t)$, and thus $\Delta \approx 0$. Otherwise, if the two states belong to different phases, and the Uhlmann factor is nontrivial, we have $\Delta \neq 0$. Thus, sudden departure of Δ from zero (for $\delta t \ll 1$) signals the points of PTs. Since both the Uhlmann parallel transport and the fidelity give rise to the same metric (the Bures metric), the nonanalyticity of Δ is accompanied by the same behavior of the fidelity. Note that the other way around is not necessarily true: in case the states commute with each other and differ only in their eigenvalues, the Uhlmann connection is trivial, and thus $\Delta = 0$.

In order for the Uhlmann connection and the fidelity to be in tune, they must be taken over the same base space. In previous studies [13], an Uhlmann connection in 1D translationally invariant systems was considered. The base space considered is the momentum space and the density matrices are of the form $\{\rho_k := e^{-\beta H(k)}/Z : k \in \mathcal{B}\}$, where $H(k) = E(k)\vec{n}(k) \cdot \vec{\sigma}/2$ and $\mathcal{B}$ is the first Brillouin zone. Since we are in one dimension, there exists no curvature and hence the holonomy along the momentum space cycle becomes a topological invariant (depends only on the homotopy class of the path). It was found that the Uhlmann geometric phase $\Phi_U(\gamma_c)$ along the closed curve given by $\gamma_c(k) = \rho_k$, changes abruptly from π to 0 after some "critical" temperature T_U. Namely, the Uhlmann phase is given by

$$\Phi_U(\gamma_c) = \arg \mathrm{Tr}\{w(-\pi)^\dagger w(\pi)\} = \arg \mathrm{Tr}\{\rho_\pi U(\gamma_c)\}, \quad (3)$$

where $w(k)$ is the horizontal lift of the loop of density matrices ρ_k, and $U(\gamma_c)$ is the Uhlmann holonomy along the first Brillouin zone. This temperature, though, is not necessarily related to a physical quantity that characterizes a system's phase. It might be the case that the Uhlmann phase is trivial, $\Phi_U(\gamma_c) = 0$, while the corresponding holonomy is not, $U(\gamma_c) \neq I$. For the systems studied in Ref. [13], the Uhlmann holonomy is a smooth function of the temperature and is given, in the basis in which the chiral symmetry operator is diagonal, by

$$U(\gamma_c) = \exp\left\{-\frac{i}{2}\int_{-\pi}^{\pi}\left[1 - \mathrm{sech}\left(\frac{E(k)}{2T}\right)\right]\frac{\partial\varphi}{\partial k}dk\sigma_z\right\}, \quad (4)$$

where $\varphi(k)$ is the polar angle coordinate of the vector $\vec{n}(k)$ lying on the equator of the Bloch sphere. Note that $\lim_{T\to 0} U(\gamma_c) = e^{-i\nu\pi\sigma_z}$, with the Berry phase being $\Phi_B = \lim_{T\to 0}\Phi_U = \nu\pi$, and ν the winding number. While in this case the Uhlmann phase suffers from an abrupt change (steplike behavior), the Uhlmann holonomy is smooth and there is no PT-like behavior.

In the paradigmatic case of the quantum Hall effect, at $T = 0$, the Hall conductivity is quantized in multiples of the first Chern number of a vector bundle in momentum space through several methods. For example, one can use linear response theory or integrate the fermions to obtain the effective action of an external U(1) gauge field. The topology of the bands appears, thus, in the response of the system to an external field. It is unclear, though, that the former mathematical object, the Uhlmann geometric phase along the cycle of the 1D momentum space, has an interpretation in terms of the response of the system. In order to measure this Uhlmann geometric phase, one would have to be able to change the quasimomentum of a state in an adiabatic way. In realistic setups, the states at finite temperatures are statistical mixtures over all momenta, such as the thermal states considered, and realizing closed curves of states ρ_k with precise momenta changing in an adiabatic way seems a tricky task. The fidelity computed in our Letter though, refers to the change of the system's *overall* state, with respect to its parameters (controlled in the laboratory much like an external gauge field), and is related to an, *a priori*, physically relevant geometric quantity, the Uhlmann factor U. The quantity $\Delta = \mathrm{Tr}[|\sqrt{\rho(t+\delta t)}\sqrt{\rho(t)}|(I - U)]$ contains information concerning the Uhlmann factor, since $U = U(t+\delta t)U^\dagger(t)$, where $U(t) = T\exp\{-\int_0^t \mathcal{A}(d\rho/ds)ds\}$ is the parallel transport operator and $\mathcal{A}$ is the Uhlmann connection differential 1-form (for details, see Ref. [42]).

Results.—In our analysis we probe the fidelity and Δ with respect to the parameters of the Hamiltonian describing the system and the temperature, independently. We perform this analysis for paradigmatic models of TIs (SSH and Creutz ladder) and TSCs (Kitaev Chain) in one dimension. We analytically calculate the expressions for the fidelity and Δ, for thermal states $\rho = e^{-\beta H}/Z$, where β is the inverse temperature (see the Supplemental Material, SM1 [43] for the details of the derivation). We use natural units: $\hbar = k_B = 1$.

Here we focus on the Creutz ladder model, while the results for SSH and the Kitaev chain are presented in the Supplemental Material, SM2 [43], since they are qualitatively the same. The Hamiltonian for the Creutz ladder model [36,37] is given by

$$\mathcal{H} = -\sum_{i\in\mathbb{Z}} K(e^{-i\phi}a_{i+1}^\dagger a_i + e^{i\phi}b_{i+1}^\dagger b_i) + K(b_{i+1}^\dagger a_i + a_{i+1}^\dagger b_i) + M a_i^\dagger b_i + \mathrm{H.c.}, \quad (5)$$

where a_i, b_i, with $i \in \mathbb{Z}$, are fermion annihilation operators, K and M are hopping amplitudes (horizontal or diagonal and vertical, respectively) and $e^{i\phi}$ is a phase factor associated with a discrete gauge field. We take $2K = 1$, $\phi = \pi/2$. Under these conditions, the system is topologically nontrivial when $M < 1$ and trivial when $M > 1$. Given two close points (M, T) and $(M', T') = (M + \delta M, T + \delta T)$, we compute $F(\rho, \rho')$ and $\Delta(\rho, \rho')$ between the states $\rho = \rho(M, T)$ and $\rho' = \rho(M', T')$. To distinguish the contributions due to the change of the Hamiltonian's parameter and the temperature, we consider the cases $\delta T = 0$ and $\delta M = 0$, respectively; see Fig. 1.

We see that for $T = 0$ both fidelities exhibit a sudden drop in the neighborhood of the gap-closing point $M = 1$, signaling the topological quantum PT. As temperature increases, the drops of both fidelities at the quantum critical point are rapidly smoothened towards the $F = 1$ value. This shows the absence of both finite-temperature parameter-driven, as well as temperature-driven (i.e., thermal) PTs. The plot for Δ, for the case $\delta T = 0$, shows a behavior similar to that of the fidelity, while for $\delta M = 0$ we obtain no information, as Δ is identically equal to zero, due to the triviality of the Uhlmann connection associated with the mutually commuting states (a consequence of the Hamiltonian's independence on the temperature). Δ is sensitive to PTs for which the state change is accompanied by a change of the eigenbasis (in contrast to fidelity, which is sensitive to both changes of eigenvalues and eigenvectors). For TIs and TSCs, this corresponds to parameter-driven transitions only.

We further study a topologically trivial superconducting system, given by the BCS theory, with the effective Hamiltonian

$$\mathcal{H} = \sum_k (\varepsilon_k - \mu)c_k^\dagger c_k - \Delta_k c_k^\dagger c_{-k}^\dagger + \mathrm{H.c.}, \quad (6)$$

where ε_k is the energy spectrum, μ is the chemical potential, Δ_k is the superconducting gap, $c_k \equiv c_{k\uparrow}$ and $c_{-k} \equiv c_{-k\downarrow}$ are operators annihilating, respectively, an electron with

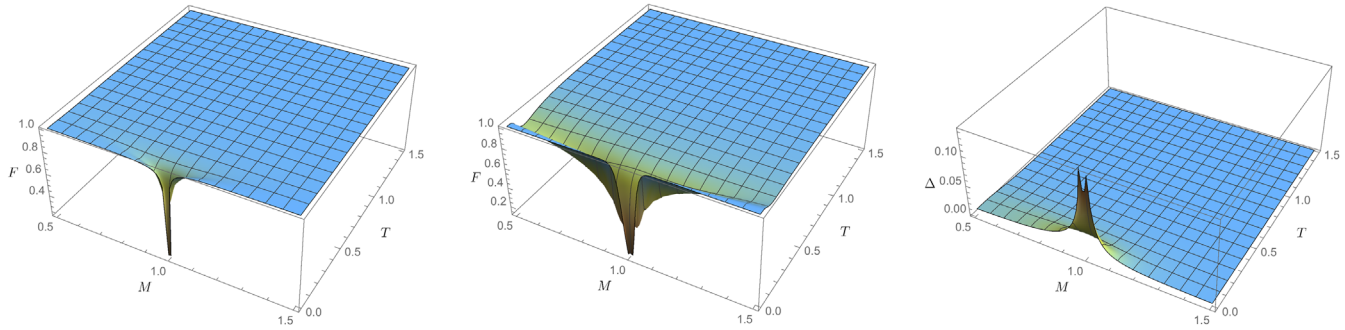


FIG. 1. The fidelity for thermal states ρ, when probing the parameter of the Hamiltonian that drives the topological PT $\delta M = M' - M = 0.01$ (left), and the temperature $\delta T = T' - T = 0.01$ (center), and the Uhlmann connection, when probing the parameter of the Hamiltonian M (right), for the Creutz ladder model (representative of the symmetry class AIII). The plot for Δ when deforming the thermal state along T is omitted since it is equal to zero everywhere.

momentum k and spin-up and an electron with momentum $-k$ and spin-down. The gap parameter is determined in the above mean-field Hamiltonian through a self-consistent mass gap equation and it depends on the original Hamiltonian's coupling associated with the lattice-mediated pairing interaction V, absorbed in Δ_k (for more details, see Ref. [16]). The solution of the equation renders the gap temperature dependent. In Fig. 2, we show the quantitative results for the fidelity and Δ. We observe that both quantities show the existence of thermally driven PTs, as their abrupt change in the point of criticality at $T = 0$, survive and drift, as temperature increases. Unlike TSCs, in this model the temperature does not only appear in the thermal state, but it is also a parameter of the effective Hamiltonian, resulting in the change of the system's eigenbasis and, consequently, nontrivial Uhlmann connection. For a detailed analysis and the explanation of the differences between the two, see the Supplemental Material, SM4 [43].

Finally, we also studied the behavior of the edge states for the TIs and the Majorana modes for the Kitaev chain, on open chains of 500 and 300 sites, respectively. In the case of TIs, we showed that the edge states localized at the boundary between two distinct topological phases, present at zero temperature, are gradually smeared out with the temperature increase, confirming the absence of temperature-driven PTs (see the Supplemental Material, SM3.1 [43] for detailed quantitative results and technical analysis). Our results on the edge states, obtained for systems in thermal equilibrium, agree with those concerning open systems treated within the Lindbladian approach [10] (and, consequently, due to considerable computational hardness, obtained for an open chain of 8 sites). Similarly, we showed that the Majorana modes exhibit an abrupt change at the zero-temperature point of the quantum PT, while the finite-temperature behavior is smooth, confirming the results obtained through the fidelity and the Uhlmann connection analysis (see the Supplemental Material, SM3.2 [43]). In the case of zero-temperature, Majorana modes of the Kitaev model are known to be good candidates for encoding qubit states in stable quantum memories, see Refs. [50,51] and references therein. The presence of robust Majorana modes at low, but finite, temperatures is a significant property which could be used in designing stable quantum memories in realistic setups.

Conclusions and outlook.—We studied the relationship between the fidelity and the Uhlmann connection over the system's *parameter space* (including parameters of the system's Hamiltonian and temperature) and found that their behaviors are consistent whenever the variations of the parameters produce variations in the eigenbasis of the density matrix. By means of this analysis, we showed the absence of temperature-driven PTs in 1D TIs and TSCs. We clarified that the Uhlmann geometric phase considered in *momentum space* is not adequate to infer such PTs, since it is only a part of the information contained in the Uhlmann holonomy. Indeed, this holonomy, as a function of temperature, is smooth [Eq. (4)]; hence, no PT-like phenomenon is expected. Furthermore, we performed the same analysis in the case of BCS superconductivity, where, in contrast to the former

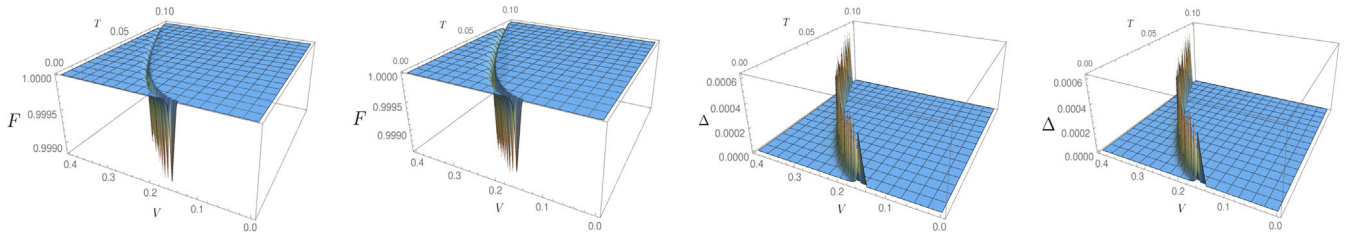


FIG. 2. The fidelity for thermal states ρ when probing the parameter of the Hamiltonian $\delta V = V' - V = 10^{-3}$ (left) and the temperature $\delta T = T' - T = 10^{-3}$ (center left), and the Uhlmann connection (center right and right, respectively), for BCS superconductivity.

systems, thermally driven PTs occur and are captured by both the fidelity and the Uhlmann connection. This shows that, when changing the temperature, the density operator is changing both at the level of its spectrum and its eigenvectors. We analyzed in detail the origin of the differences between the BCS and the Kitaev chain and suggested that in realistic scenarios the gap of topological superconductors could also, generically, be temperature dependent.

Finally, we would like to point out possible future lines of research. The study of Majorana modes at finite temperature suggests that they can be used in achieving realistic quantum memories. This provides a relevant path for future research. Another related subject is to perform the same analysis using an open system approach where the system interacts with a bath and eventually thermalizes. There the parameter space would also include the parameters associated with the system-bath interaction.

B. M. and C. V. acknowledge the support from DP-PMI and FCT (Portugal) through the Grants No. SFRH/BD/52244/2013 and No. PD/BD/52652/2014, respectively. B. M. acknowledges the support of the Physics of Information and Quantum Technologies Group. C. V. and N. P. acknowledge the support of SQIG—Security and Quantum Information Group and UID/EEA/50008/2013. N. P. acknowledges the IT project QbigD funded by FCT PEst-OE/EEI/LA0008/2013. Support from FCT (Portugal) through Grant No. UID/CTM/04540/2013 is also acknowledged.

[1] P. W. Anderson, *Basic Notions of Condensed Matter Physics* (Westview Press, Boulder, CO, 1997).
[2] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
[3] M. V. Berry, Proc. R. Soc. A **392**, 45 (1984).
[4] B. Zeng, X. Chen, D.-L. Zhou, and X.-G.Wen, arXiv:1508.02595.
[5] Y. Ando, J. Phys. Soc. Jpn. **82**, 102001 (2013).
[6] A. Kitaev, AIP Conf. Proc. **1134**, 22 (2009).
[7] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, Phys. Rev. B **78**, 195125 (2008).
[8] S. Ryu and Y. Hatsugai, Phys. Rev. Lett. **89**, 077002 (2002).
[9] A. Rivas, O. Viyuela, and M. A. Martin-Delgado, Phys. Rev. B **88**, 155141 (2013).
[10] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Phys. Rev. B **86**, 155140 (2012).
[11] A. Uhlmann, Ann. Phys. (N.Y.) **501**, 63 (1989).
[12] M. Hübner, Phys. Lett. A **179**, 226 (1993).
[13] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Phys. Rev. Lett. **112**, 130401 (2014).
[14] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Phys. Rev. Lett. **113**, 076408 (2014).
[15] Z. Huang and D. P. Arovas, Phys. Rev. Lett. **113**, 076407 (2014).
[16] N. Paunković and V. R. Vieira, Phys. Rev. E **77**, 011129 (2008).
[17] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, 2D Mater. **2**, 034006 (2015).
[18] O. Andersson, I. Bengtsson, M. Ericsson, and E. Sjöqvist, Phil. Trans. R. Soc. A **374**, 20150231 (2016).
[19] S. N. Kempkes, A. Quelle, and C. Morais Smith, Sci. Rep. **6**, 38530 (2016).
[20] J. C. Budich and S. Diehl, Phys. Rev. B **91**, 165140 (2015).
[21] J. Tidström and E. Sjöqvist, Phys. Rev. A **67**, 032110 (2003).
[22] J. Åberg, D. Kult, E. Sjöqvist, and D. K. L. Oi, Phys. Rev. A **75**, 032106 (2007).
[23] O. Viyuela, A. Rivas, S. Gasparinetti, A. Wallraff, S. Filipp, and M. Martin-Delgado, A measurement protocol for the topological Uhlmann phase, arXiv:1607.08778v1.
[24] J. Zhu, M. Shi, V. Vedral, X. Peng, D. Suter, and J. Du, Europhys. Lett. **94**, 20007 (2011).
[25] A. Hamma, R. Ionicioiu, and P. Zanardi, Phys. Lett. A **337**, 22 (2005).
[26] A. Hamma, W. Zhang, S. Haas, and D. A. Lidar, Phys. Rev. B **77**, 155111 (2008).
[27] G. B. Halász and A. Hamma, Phys. Rev. A **86**, 062330 (2012).
[28] D. F. Abasto, A. Hamma, and P. Zanardi, Phys. Rev. A **78**, 010301 (2008).
[29] P. Zanardi and N. Paunković, Phys. Rev. E **74**, 031123 (2006).
[30] J.-H. Zhao and H.-Q. Zhou, Phys. Rev. B **80**, 014403 (2009).
[31] T. P. Oliveira and P. D. Sacramento, Phys. Rev. B **89**, 094512 (2014).
[32] N. Paunković, P. D. Sacramento, P. Nogueira, V. R. Vieira, and V. K. Dugaev, Phys. Rev. A **77**, 052302 (2008).
[33] P. Zanardi, L. Campos Venuti, and P. Giorda, Phys. Rev. A **76**, 062318 (2007).
[34] A. C. M. Carollo and J. K. Pachos, Phys. Rev. Lett. **95**, 157203 (2005).
[35] M. E. Reuter, M. J. Hartmann, and M. B. Plenio, Proc. Math. Phys. Eng. Sci. **463**, 1271 (2007).
[36] M. Creutz, Phys. Rev. Lett. **83**, 2636 (1999).
[37] A. Bermudez, D. Patané, L. Amico, and M. A. Martin-Delgado, Phys. Rev. Lett. **102**, 135702 (2009).
[38] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).
[39] A. Y. Kitaev, Phys. Usp. **44**, 131 (2001).
[40] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
[41] A. Uhlmann, Found. Phys. **41**, 288 (2011).
[42] D. Chruscinski and A. Jamiolkowski, *Geometric Phases in Classical and Quantum Mechanics* (Springer Science & Business Media, New York, 2012), Vol. 36.
[43] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.119.015702 for references on the discontinuity development in the thermodynamic limit [44], and on the use of quantum memories in cryptography, including Refs. [45–49].
[44] H. T. Quan and W. H. Zurek, New J. Phys. **12**, 093025 (2010).
[45] J. Rodrigues, P. Mateus, N. Paunković, and A. Souto, J. Phys. A **50**, 205301 (2017).
[46] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Gehring, C. S. Jacobsen, and U. L. Andersen, Nat. Photonics **9**, 397 (2015).

[47] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, Nat. Phys. **6**, 659 (2010).
[48] R. Loura, A. J. Almeida, P. S. André, A. N. Pinto, P. Mateus, and N. Paunković, Phys. Rev. A **89**, 052336 (2014).
[49] R. Loura, D. Arsenović, N. Paunković, D. B. Popović, and S. Prvanović, Phys. Rev. A **94**, 062335 (2016).
[50] J. Alicea, Rep. Prog. Phys. **75**, 076501 (2012).
[51] M. Ippoliti, M. Rizzi, V. Giovannetti, and L. Mazza, Phys. Rev. A **93**, 062325 (2016).

Reply to "Comment on 'Bit-string oblivious transfer based on quantum state computational distinguishability'"

A. Souto,[1,2,*] P. Mateus,[1,2,†] P. Adão,[1,3,‡] and N. Paunković[1,2,§]

[1]*SQIG—Instituto de Telecomunicações, Avenida Rovisco Pais 1, Torre Norte, Piso 10, PT-1049-001 Lisbon, Portugal*

[2]*Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, PT-1049-001 Lisbon, Portugal*

[3]*Departamento de Engenharia Informática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal*

(Received 24 July 2015; published 30 October 2015)

In the Comment the author states that the proposed all-or-nothing oblivious transfer (OT) protocol in our paper is insecure against a dishonest Alice and, as a corollary, derives an attack to Crépeau's construction of 1-out-of-2 OT. The security criterion used in the Comment is indeed stronger than the one used in our paper. However, we argue that the criterion used in our paper is in the spirit of the original idea of the OT protocol proposed by Rabin. Moreover, a protocol that satisfies the criterion in our paper can be used to construct useful multiparty protocols. Finally, the protocol in our paper can be used, together with a secure bit commitment scheme, to construct a 1-out-of-2 OT secure against malicious Alice, achieving the security requirement considered in the Comment.

DOI: 10.1103/PhysRevA.92.046302 PACS number(s): 03.67.Dd, 03.67.Ac

In the Comment [1], the author suggests that our protocol presented in our paper [2] fails on two points: (i) It does not meet the security requirement of the all-or-nothing oblivious transfer (OT) protocol, and (ii) it is unsuitable for serving as a building block for 1-out-of-2 oblivious transfer in the reduction scheme proposed by Crépeau [3]. In fact, our protocol is secure against the criterion defined in our paper, and while the security criterion proposed by the author is indeed a sensible one, we argue below that our definition is also both in the spirit of the original idea of the Rabin OT [4] and is a suitable one for applications (see, for example, Refs. [5,6]).Regarding (ii), although one cannot achieve 1-out-of-2 OT secure against malicious Alice using Crépeau's reduction [3] (as suggested in the Comment), it is possible to do so following the ideas presented in Ref. [7].

(i) In the proposed attack, a cheating Alice sends a wrong "key" (wrong permutation π), and, as a consequence, Bob obviously does not get any message. This is, however, the same as not sending the message at all, i.e., sending a random message, say, a completely mixed state. In the original idea of Rabin [4], and also expressed in Refs. [8,9], the goal of Alice in OT protocols is to reveal secrets to Bob (the title of Rabin's paper is in fact "How to exchange secrets by oblivious transfer"). Moreover, "*Rabin OT in fact corresponds to a binary erasure channel with erasure probability 1/2*" [10], a channel in which Alice *does send* a message to Bob, which then might be either lost or delivered, without Alice being aware of what happened.

In the proposed attack, Alice is not giving Bob a chance to open the secret intended to be shared. If Alice does not want to share a secret, then there is not much one can do to prevent it and in that case she obviously knows that Bob did not receive it. But this cheating strategy does not fall within the scope of the OT protocol discussed in our paper, where it is assumed that a successful cheating strategy is one in which *only one* security criterion is violated, while the others are satisfied (a cheating Alice does want to know if Bob got the message, but still wants to give him a 50% chance of opening the message).

(ii) The proposed attack [1] to Crépeau's reduction scheme is indeed correct, but a *dishonest* Alice can perform a similar attack using any all-or-nothing OT that is secure according to the stronger criterion considered in the Comment [1]. Namely, she can simply [in Step (7) of the protocol below] flip a fair coin and, depending on its outcome, either proceed with executing Crépeau's reduction honestly, or send the same value in both messages (bits):

Protocol 1 (Cheating 1*-out-of-*2 *OT)*

(1) Alice and Bob agree on a security parameter s.

(2) Alice chooses at random Ks bits $r_1, r_2, \ldots, r_{Ks}$ for some constant K.

(3) Alice runs the secure all-or-nothing OT Ks times.

(4) Bob selects $U = \{i_1, i_2, \ldots, i_{\alpha_s}\}$ and $V = \{i_{\alpha_s+1}, i_{\alpha_s+2}, \ldots, i_{2\alpha_s}\}$, where $\alpha_s = \lceil \frac{Ks}{3} \rceil$ with $U \cap V = \emptyset$ and such that he knows r_{i_ℓ} for each $i_\ell \in U$.

(5) Bob sends $(X,Y) = (U,V)$ or $(X,Y) = (V,U)$ to Alice according to a random bit k.

(6) Alice computes $m_0 = \bigoplus_{x \in X} r_x$ and $m_1 = \bigoplus_{y \in Y} r_y$.

(7) Alice flips a fair coin and if the outcome is heads she chooses an index $i_c \notin X \cup Y$ and continues with Crépeau's protocol, that is, sends to Bob $b_j \oplus m_0$ and $b_{\bar{j}} \oplus m_1$ for a random bit j; otherwise, she chooses $i_c \in X \cup Y$ and deviates from the protocol by sending to Bob $b_j \oplus m_0$ and $b_j \oplus m_1$, using the *same* bit b_j for a random bit j;

(8) Bob computes $\bigoplus_{u \in U} r_u \in \{m_0, m_1\}$ and uses it to get his secret bit b_k.

The above protocol has the same behavior, with respect to Alice's knowledge of the bit received by Bob, as the one proposed in the Comment [1]: In 50% of the cases, Alice knows "*which of the secret bits b_0 or b_1 is finally known by Bob.*" A cheating Alice can always choose not to send the intended messages [Step (7)] and there is nothing that Bob can do to prevent that behavior. Moreover, in our Protocol 1, Alice can choose which bit value Bob will get, while in the "Protocol for 1-out-of-2 OT" from the Comment [1] Alice can only learn

*a.souto@math.ist.utl.pt
†pmat@math.ist.utl.pt
‡pedro.adao@ist.utl.pt
§npaunkovic@math.ist.utl.pt

 ©2015 American Physical Society

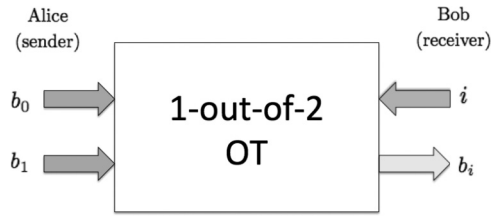


FIG. 1. Alice sends two bits b_0 and b_1 labeled by 0 and 1, respectively. Bob inputs bit $i \in \{0,1\}$, choosing to receive b_i as the output of the 1-out-of-2 OT protocol.

which of the two bit values Bob got. We see that both our all-or-nothing OT [2] and the 1-out-of-2 OT presented in the Comment (and in Ref. [3]) fail to satisfy the security criterion against malicious Alice for essentially the same reason: the impossibility to enforce Alice to send valid message(s).

On the other hand, one can still consider the so-called semihonest scenario, in which Alice follows the protocol, but is allowed to perform additional computations using the data gathered while running the protocol (for details, see, for example, Ref. [11], Chap. 7, or Ref. [12], Chap. 13). In the context of a semihonest Alice, the 1-out-of-2 OT based on our all-or-nothing OT (through the mentioned Crépeau's reduction [3]) can be used to achieve secure multiparty computation, either based on Yao's garbled circuits [5] (cited in the original paper as one of the main motivations for considering OT as a cryptographic primitive), or through the construction proposed in Ref. [6]. Assuming a semihonest Alice in the mentioned constructions is indeed a reasonable one. For example, considering a semihonest model is enough for hospitals that wish to share information. Also, the construction achieving multiparty computation proposed in Ref. [6] assumes that the majority of the agents are not malicious, which is indeed often the case (the majority of the population are honest people).

An all-or-nothing OT, secure against the criterion adopted in the Comment [1], can nevertheless be used to, following the same reduction [3], construct a stronger version of the 1-out-of-2 OT which is secure against malicious Alice. By stronger we mean that, instead of a random choice of received bit value b_i, with $i \in \{0,1\}$, it is Bob who decides which of the two bits to receive—he chooses $i \in \{0,1\}$—while Alice is oblivious of this choice, and not just the value b_i of the received bit (a notion of 1-out-of-2 OT that is predominantly used in recent literature [13–17]). For details, see Fig. 1.

While the above construction does not work with our protocol, one can construct a 1-out-of-2 OT secure against malicious Alice using a semihonest 1-out-of-2 OT (based on our all-or-nothing OT), together with a bit commitment scheme, as shown in Ref. [7].

In the following, we present the details of the reduction that uses our protocol as a black box, together with a bit commitment scheme, to construct a 1-out-of-2 OT that is secure against a malicious Alice (which includes the attack proposed in the Comment [1]). We strictly follow the idea presented in Ref. [7], which is a development over Refs. [18,19].

First, notice that our all-or-nothing OT protocol is secure against semihonest Alice, and against malicious Bob. Therefore, the 1-out-of-2 OT obtained by Crepéau's construction [3] that uses our all-or-nothing OT protocol as a black box is consequently secure against semihonest Alice and malicious Bob. From this protocol, using the OT-reversal scheme introduced in Ref. [20], one can construct a 1-out-of-2 OT protocol that is secure against malicious Alice (the sender) and semihonest Bob (the receiver). Below, we schematically recall the OT-reversal protocol:

Protocol 2 (OT-Reversal)

Alice's input: two bits s_0 and s_1.

Bob's input: a bit r.

Black box: A 1-out-of-2 OT protocol secure against semihonest sender and malicious receiver.

(1) Bob chooses bit ρ and computes $\rho \oplus r$. He performs the 1-out-of-2 OT to obliviously transfer bits b_0 and b_1, choosing either $b_0 = \rho$ and $b_1 = \rho \oplus r$, or $b_0 = \rho \oplus r$ and $b_1 = \rho$.

(2) Alice opens the bit $a = b_{s_0 \oplus s_1}$ and sends to Bob $\alpha = s_0 \oplus a$.

(3) Using α Bob computes the output $\rho \oplus \alpha$.

It is easy to see that Bob gets the desired bit s_r, i.e., $s_r = \rho \oplus \alpha$. Assuming, without the loss of generality, that Bob's choice was $b_0 = \rho$ and $b_1 = \rho \oplus r$, we have that:

(i) If $s_0 \oplus s_1 = 0$, then $s_0 = s_1$, and therefore Alice opens bit $a = b_0 = \rho$. She then sends to Bob bit $\alpha = s_0 \oplus a = s_0 \oplus \rho$. As a result, Bob's output is $\rho \oplus \alpha = \rho \oplus (s_0 \oplus \rho) = s_0 = s_1$.

(ii) Otherwise, $s_0 \oplus s_1 = 1$ and thus $s_0 = \neg s_1$. Therefore, Alice opens bit $a = b_1 = \rho \oplus r$ and sends to Bob bit $\alpha = s_0 \oplus a = s_0 \oplus (\rho \oplus r)$. As a result, Bob's output is $\rho \oplus \alpha = \rho \oplus [s_0 \oplus (\rho \oplus r)] = s_0 \oplus r$. In case $r = 0$, we have that Bob's output is $s_0 \oplus r = s_0$. Otherwise, if $r = 1$, the output is $s_0 \oplus r = \neg s_0 = s_1$.

We have started with a 1-out-of-2 OT secure against a semihonest Alice (the sender) and malicious Bob (the receiver). Using the above OT-Reversal Protocol 2, we have turned it into 1-out-of-2 OT secure against malicious Alice and semihonest Bob. To finally achieve 1-out-of-2 OT secure against both malicious agents, we need to improve the security against Bob (the receiver). This is done by the construction introduced in Ref. [7], which, in addition to semihonest 1-out-of-2 OT, uses a bit commitment scheme as well (note that this construction improves the receiver's security level regardless of the sender's behavior, i.e., whether Alice follows the protocol or she is behaves maliciously). Below, we present in detail the construction from Ref. [7].

Protocol 3 (Comp)

Alice's input: two bit strings s_0 and s_1.

Bob's input: a random value r.

Black boxes: OT protocol secure against semihonest receivers and a secure bit commitment scheme.

Phase I: Random tape generation

1. Bob chooses $2n$ pairs $(r_1^B, \tau_1^B), \ldots, (r_{2n}^B, \tau_{2n}^B)$ of random bits $r_i^B \in \{0,1\}$ and strings $\tau_i^B \in \{0,1\}^\ell$, and using a secure bit commitment as a black box, commits to those strings that are sent to Alice.[1]

[1]It is worthwhile mentioning that Bob is committing to a random τ that he will use in each execution of the given OT protocol used as a black box. In the case of our original OT protocol from Ref. [2], τ is in fact the permutation that Bob sends to Alice.

2. Alice generates $2n$ random strings $(r_1^A,\tau_1^A),\ldots,(r_{2n}^A,\tau_{2n}^A)$, where each $r_i^A \in \{0,1\}$ and $\tau_i^A \in \{0,1\}^\ell$, and sends them to Bob.

3. Bob sets bit $r_i = r_i^B \oplus r_i^A$ and string $\tau_i = \tau_i^B \oplus \tau_i^A$ for all $i = 1,\ldots,2n$.

Phase II: Basic execution

1. Alice chooses $2n$ pairs of random pairs of bit strings $(s_1^0,s_1^1),\ldots,(s_{2n}^0,s_{2n}^1)$ as inputs of the OT protocol.

2. Alice and Bob engage in $2n$ executions of the OT protocol. In the ith execution, Alice sends (s_i^0,s_i^1) and Bob chooses to open the message with index $r_i \in \{0,1\}$, obtaining the string $s_i^{r_i}$.

Phase III: Cut-and-choose

1. Alice chooses a random $q = (q_1,\ldots,q_n) \in \{0,1\}^n$. The string q is used to define a set of indices $Q \subset \{1,\ldots,2n\}$, with $|Q| = n$, such that $Q = \{2j - q_j | j = 1,\ldots,n\}$.

2. For every $i \in Q$, Bob opens the commitment (r_i^B,τ_i^B) to Alice.

3. Alice checks the validity of each commitment by checking the consistency of τ_i^B with Bob's message obtained in the execution of the OT protocol. If it is not consistent, Alice aborts and halts the protocol.

Phase IV: Combiner

1. For every $j \notin Q$, Bob computes $\alpha_j = r \oplus r_j$ and sends $\{\alpha_j\}_{j\notin Q}$ to Alice.

2. Alice computes $\sigma_0 = s_0 \oplus (\bigoplus_{j\notin Q} s_j^{\alpha_j})$ and $\sigma_1 = s_1 \oplus (\bigoplus_{j\notin Q} s_j^{1-\alpha_j})$. She sends (σ_0,σ_1) to Bob.

3. Bob computes the outputs $s_r = \sigma_r \oplus (\bigoplus_{j\notin Q} s_j^{r_j})$.

The above reduction relies on the use of secure bit commitment protocol. There exist a number of practical quantum bit commitment schemes that, providing certain technological limitations such as imperfect and/or bounded quantum memories, achieve secure bit commitment [16,21–25]. Note that in the case of imperfect memories, no matter how small the (finite) noise is, the security achieved is based on the laws of physics, which is stronger than the computational security of the classical counterparts. Unlike the discrete states of classical bits, the states of quantum bits are continuous, therefore it is a viable assumption that, at least in the very long foreseeable future, no perfect quantum memories will exist (if ever). Moreover, there exist unconditionally secure bit commitment schemes, both classical [26] and quantum [21,27], that use relativistic effects.

Finally, one could, instead of bit commitment, use a bit-string commitment scheme, such as the one proposed by Kent [27]. In this case, one should take r_i^B to be strings instead of bits, and use the inner product (modulo 2) instead of XOR when defining bits r_i in point 3 of Phase I. Note though that the equivalence between single-bit and genuine bit-string protocols, which exists in the classical realm, does not necessarily have to be valid in the quantum setting, which is the reason why Kent's result on bit-string commitment does not contradict the bit-commitment no-go theorems [28–31]. Indeed, while no bit from a committed string in Kent's protocol could be changed with a non-negligible probability, this does not mean that a whole string cannot be altered with a non-negligible probability. The analysis of the degree of security of such implementation goes beyond the scope of this Reply.

OT protocols are security primitives, and as such, not useful *per se* (unless their application is outside the field of cryptography, for example, when modeling the binary erasure channel). Only when used (as black boxes) in more complex reduction schemes can they serve the purpose in achieving relevant real-life security protocols. The above discussion suggests that, instead of talking of "(im)proper" security criteria in absolute terms, one should perhaps rather talk of "suitable" criteria *with respect to* certain applications and their requirements. While the attack proposed in the Comment [1] is indeed valid for the case of a "unverified" all-or-nothing OT,[2] our protocol can, (i) in the context of a semihonest model, be used to achieve useful secure multiparty computations, and, (ii) together with a commitment scheme, achieve 1-out-of-2 OT secure against malicious Alice, a protocol that satisfies the stronger security criterion adopted in the Comment [1]. Finally, we note that this analysis applies to any all-or-nothing OT that fulfills our security criteria, in particular, the one presented in Refs. [32,33].

We thank Jeroen van de Graaf for very helpful discussions regarding oblivious transfer. This work was partially supported, under the CV-Quantum internal project at IT, by Fundação para a Ciência e Tecnologia (FCT) PEst-OE/EEI/LA0008/2013 and UID/EEA/50008/2013, namely, via the FCT CaPri initiative. A.S. also acknowledges the FCT postdoc Grant No. SFRH/BPD/76231/2011.

[2]By "unverified" we mean that it is not required for Bob to be able to verify if Alice indeed sent the message, as it is assumed that it is in her interest to do so, i.e., Alice is a semihonest agent.

[1] G. He, Phys. Rev. A **92**, 046301 (2015).

[2] A. Souto, P. Mateus, P. Adão, and N. Paunković, Phys. Rev. A **91**, 042306 (2015).

[3] C. Crépeau, in *Advances in Cryptology: CRYPTO '87*, edited by C. Pomerance, Lecture Notes in Computer Science Vol. 293 (Springer, Berlin, 1988), pp. 350–354.

[4] M. Rabin, Aiken Computation Laboratory, Harvard University Technical Memo No. TR-81, 1981 (unpublished).

[5] A. Yao, in *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science, SFCS '82* (IEEE Computer Society, Washington, DC, 1982), pp. 160–164.

[6] O. Goldreich, S. Micali, and A. Wigderson, in *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing, STOC '87* (ACM, New York, 1987), pp. 218–229.

[7] S. Choi, D. Dachman-Soled, T. Malkin, and H. Wee, in *Theory of Cryptography*, edited by O. Reingold, Lecture Notes in Computer Science Vol. 5444 (Springer, Berlin, 2009), pp. 387–402.

[8] C. Crépeau and J. Kilian, in *29th Annual Symposium on Foundations of Computer Science* (IEEE, New York, 1988), pp. 42–52.

[9] C. Crépeau and J. Kilian, in *Advances in Cryptology: CRYPTO '88*, edited by S. Goldwasser, Lecture Notes in Computer Science Vol. 503 (Springer, New York, 1990), pp. 2–7.

[10] C. Crépeau, K. Morozov, and S. Wolf, in *Security in Communication Networks*, edited by C. Blundo and S. Cimato, Lecture Notes in Computer Science Vol. 3352 (Springer, Berlin, 2005), pp. 47–59.

[11] O. Goldreich, *Foundations of Cryptography: Volume 2, Basic Applications* (Cambridge University Press, New York, 2004).

[12] Y. Lindell, *Foundations of Cryptography 89-856* (2010), Lecture notes written for a graduate course in the foundations of cryptography at Bar-Ilan University, Israel. Available at http://u.cs.biu.ac.il/∼lindell/89-856/complete-89-856.pdf.

[13] C. Crépeau, J. van de Graaf, and A. Tapp, in *Advances in Cryptology*, edited by D. Coppersmith, Lecture Notes in Computer Science 963 (Springer, Berlin, 1995), pp. 110–123.

[14] G. Brassard, C. Crépeau, and S. Wolf, J. Cryptology **16**, 219 (2003).

[15] C. Crépeau and G. Savvides, in *Advances in Cryptology—EUROCRYPT 2006*, edited by S. Vaudenay, Lecture Notes in Computer Science Vol. 4004 (Springer, Berlin, 2006), pp. 201–221.

[16] S. Wehner, C. Schaffner, and B. M. Terhal, Phys. Rev. Lett. **100**, 220502 (2008).

[17] C. Erven, N. Ng, N. Gigov, R. Laflamme, S. Wehner, and G. Weihs, Nat. Commun. **5**, 3418 (2014).

[18] Y. Ishai, E. Kushilevitz, Y. Lindell, and E. Petrank, in *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing, STOC '06* (ACM, New York, 2006), pp. 99–108.

[19] I. Haitner, in *Proceedings of the 5th Conference on Theory of Cryptography, TCC '08* (Springer, Berlin, 2008), pp. 412–426.

[20] S. Wolf and J. Wullschleger, in *Advances in Cryptology—EUROCRYPT 2006* (Ref. [15]), pp. 222–232.

[21] N. Bouman, S. Fehr, C. Gonzalez-Guillen, and C. Schaffner, in *Theory of Quantum Computation, Communication, and Cryptography*, edited by K. Iwama, Y. Kawano, and M. Murao, Lecture Notes in Computer Science Vol. 7582 (Springer, Berlin, 2013), pp. 29–44.

[22] C. Schaffner, B. M. Terhal, and S. Wehner, Quantum Inf. Comput. **9**, 963 (2011).

[23] R. Koenig, S. Wehner, and J. Wullschleger, IEEE Trans. Inf. Theory **58**, 1962 (2012).

[24] N. Ng, S. Joshi, C. Ming, C. Kurtsiefer, and S. Wehner, Nat. Commun. **3**, 1326 (2012).

[25] R. Loura, A. J. Almeida, P. S. André, A. N. Pinto, P. Mateus, and N. Paunković, Phys. Rev. A **89**, 052336 (2014).

[26] A. Kent, J. Cryptology **18**, 313 (2005).

[27] A. Kent, Phys. Rev. Lett. **83**, 1447 (1999).

[28] H. K. Lo and H. F. Chau, Phys. Rev. Lett. **78**, 3410 (1997).

[29] D. Mayers, Phys. Rev. Lett. **78**, 3414 (1997).

[30] G. M. D'Ariano, D. Kretschmann, D. Schlingemann, and R. F. Werner, Phys. Rev. A **76**, 032328 (2007).

[31] L. Magnin, F. Magniez, A. Leverrier, and N. J. Cerf, Phys. Rev. A **81**, 010302 (2010).

[32] J. Rodrigues, P. Mateus, N. Paunković, and A. Souto, arXiv:1407.8146.

[33] P. Mateus, N. Paunković, J. Rodrigues, and A. Souto, in *Communications and Multimedia Security—Proceedings of the 15th IFIP TC 6/TC 11 International Conference, CMS 2014, Aveiro, Portugal, September 25–26, 2014*, edited by B. De Decker and A. Zúquete, Lecture Notes in Computer Science Vol. 8735 (Springer, Berlin, 2014), pp. 147–153.

Bit-string oblivious transfer based on quantum state computational distinguishability

A. Souto,[1,2,*] P. Mateus,[1,2,†] P. Adão,[1,3,‡] and N. Paunković[1,2,§]

[1]*SQIG—Instituto de Telecomunicações, Avenida Rovisco Pais 1, Torre Norte, Piso 10, PT-1049-001 Lisbon, Portugal*

[2]*Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, PT-1049-001 Lisbon, Portugal*

[3]*Departamento de Engenharia Informática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, PT-1049-001 Lisbon, Portugal*

(Received 29 April 2014; revised manuscript received 28 October 2014; published 8 April 2015)

Oblivious transfer protocol is a basic building block in cryptography and is used to transfer information from a sender to a receiver in such a way that, at the end of the protocol, the sender does not know if the receiver got the message or not. Since Shor's quantum algorithm appeared, the security of most of the classical cryptographic schemes has been compromised, as they rely on the fact that factoring is infeasible. To overcome this, quantum mechanics has been used intensively in the past decades, and alternatives resistant to quantum attacks have been developed in order to fulfill the (potential) lack of security of a significant number of classical schemes. In this paper, we present a quantum computationally secure protocol for bit-string oblivious transfer between two parties, under the assumption of quantum hardness of state distinguishability and the constraint of performing at most few-qubit measurements (leaving open the question of general attacks performed on all qubits involved). The protocol is feasible, in the sense that it is implementable in polynomial time.

DOI: 10.1103/PhysRevA.91.042306 PACS number(s): 03.67.Dd, 03.67.Ac

I. INTRODUCTION

An *oblivious transfer* protocol involves two parties, a sender (Alice) and a receiver (Bob). It consists of two phases: the transferring phase and the opening phase. The goal of the sender is to send a message (in general, a bit string) during the transferring phase, that will not be known to the receiver until the opening phase, during which the sender reveals the message with probability $1/2$. The goal of the receiver is that, upon opening the message, the sender is oblivious to the fact whether the message was successfully transferred or not. Although not explicitly stated in the original argument, it is usually assumed that the receiver knows, at the end of the protocol, if he got the correct message (see, for example, [1]).

In classical cryptography, the first oblivious transfer scheme was proposed by Rabin [2] and is based on the same assumptions as the RSA cryptographic system. In Rabin's scheme, the sender sends a message to the receiver that is able to decrypt it properly with probability $1/2$. At the end of the protocol, the sender remains oblivious to whether or not the receiver got the correct message. Later, Even, Goldreich, and Lempel [3] formally presented the idea of *1-out-of-2 oblivious transfer*. In 1-out-of-2 oblivious transfer the sender has two messages to send such that the receiver gets only one of the two with equal probability; the sender is oblivious to which message was received. It is clear that from a 1-out-of-2 oblivious transfer protocol, by choosing one of the messages at random, it is possible to construct an oblivious transfer protocol. Furthermore, it is possible to reduce the 1-out-of-2 oblivious transfer protocol, where the message is a bit string, to a single-bit 1-out-of-2 oblivious transfer protocol, first shown in [4] (see also [5,6]). Finally, Crépeau showed that Rabin's oblivious transfer is equivalent to 1-out-of-2 oblivious transfer [1]). This way, a single-bit oblivious transfer becomes a cryptographic primitive for all different flavors of oblivious transfer protocols, in *classical setting*. Moreover, oblivious transfer is a particularly important primitive as from $O(n)$ instances of a 1-out-of-2 oblivious transfer protocol one can construct a bit-commitment scheme, using the technique presented in Ref. [7].

In the last decades quantum computation has played a crucial role in the development of cryptographic analysis. The breakthrough of quantum computation in the realm of cryptography is due to Shor's factoring algorithm [8]. This algorithm compromises the security of most common public key cryptographic schemes as factoring becomes feasible with a quantum computer. Since today's technology is evolving to be able to deal with a larger number of qubits, a possibility of having affordable and reliable quantum computers in the future arises, compromising the secrecy of communications and transactions.

The first results in the field of quantum cryptography were made already back in 1969 by Wiesner, when he introduced the notions of quantum multiplexing and quantum money. The former was essentially a 1-out-of-2 oblivious transfer protocol which was introduced more than a decade later [3] within the scope of classical cryptography. Unfortunately, his ideas remained largely unknown to the general scientific community as he managed to publish his work only in 1983 [9]. Subsequently, based on similar ideas to those of Wiesner, Bennet and Brassard developed their famous BB84 cryptographic protocol for key distribution [10], which was shown to be unconditionally secure [11–14], in contrast to only computationally secure classical cryptography.

Following the success of quantum cryptography, a significant effort was put in developing other quantum security protocols that would outperform their classical counterparts. Unfortunately, a number of no-go results showed the impossibility of having some of the important security primitives unconditionally secure even in the quantum setting (nevertheless

*a.souto@math.ist.utl.pt
†pmat@math.ist.utl.pt
‡pedro.adao@ist.utl.pt
§npaunkovic@math.ist.utl.pt

 ©2015 American Physical Society

some, such as quantum contract signing, have been shown to be unconditionally secure [15]). First, Lo and Chau [16], and independently Mayers [17], showed the impossibility of designing an unconditionally secure quantum bit commitment protocol, unless relativistic effects are used (see [18–20]). This resulted in the development of various practically secure quantum bit commitment protocols using noisy and bounded memories [20–25].

Subsequently, Lo showed the same no-go result for a broad class of two-party security protocols, called "one-sided two-party computations" [26]. An immediate corollary to this result is the impossibility of an unconditionally secure quantum 1-out-of-2 oblivious transfer, in particular transfer of single bits. Note that in the classical setting, due to the above mentioned reduction of a bit commitment to a 1-out-of-2 oblivious transfer, the impossibility of having a secure bit commitment would imply the impossibility of having a secure single-bit 1-out-of-2 oblivious transfer as well. In the quantum realm the situation is more subtle and complicated, and classical reductions may not be valid anymore.

Namely, classical reduction schemes are typically based on a number of executions of one ("more basic") protocol, to achieve another ("more complex") one, where the "simpler" protocol is treated as a black box. In quantum mechanics such an assumption is generally not valid, as the parties involved in the protocol are in principle allowed to perform joint measurements on different "black boxes," thus compromising the classical reduction schemes. Indeed, although secure quantum single-bit 1-out-of-2 oblivious transfer was shown to be impossible, a secure quantum single-bit oblivious transfer is possible [27]. This result showed that the two "flavors" of oblivious transfer protocols are no longer equivalent in the quantum domain, as further elaborated in Ref. [28]. The reason for this is precisely in the fact that the above mentioned Crépeau's reduction scheme [1] uses many instances of oblivious transfer protocol in order to construct a 1-out-of-2 protocol, thus making it vulnerable to so-called "coherent attacks" (joint measurements). Hence, other cryptographic protocols based on multiple usage of simpler protocols (as black boxes) have to be reanalyzed when quantum counterparts are considered. In particular, the use of oblivious transfer as a building block for more complex cryptographic protocols [4,29–32], and in designing secure multiparty computation schemes, as suggested in recent papers [33,34]. A possible approach towards a rigorous security framework for complex protocols based on the above mentioned primitives could be taken along the lines presented in Refs. [35,36], where the so-called universal composability of quantum key distribution-based schemes was considered.

Furthermore, Kent showed that, although unconditionally secure quantum bit commitment is not possible, a secure quantum bit-string protocol is [37], thus showing explicitly the insecurity of classical reduction schemes from a single-bit to a bit-string protocol (see also a discussion on a slightly different version of a bit-string commitment [38], as well as a quantum bit-string generation protocol [39]). Thus, although there exists an unconditionally secure quantum single-bit oblivious transfer, a question of a quantum bit-string counterpart stays open.

In this paper we present a polynomial time quantum bit-string oblivious transfer protocol based on a presumed polynomial hard problem even for a quantum computer [40], the *Quantum State Computational Distinguishability with Fully Flipped Permutations*, denoted by QSCD_{ff}, presented in Ref. [41]. We prove the protocol's security against at most few-qubit coherent attacks, leaving open the question of most general multiqubit measurements. We propose here a concrete quantum realization of a bit-string oblivious transfer protocol based on a presumably quantum hard problem. QSCD_{ff} is the standard problem of quantum state distinguishability, applied to two particular quantum states: $\otimes_{i=1}^{k(n)} \frac{1}{\sqrt{2}}(|\sigma_i\rangle + |\sigma_i \circ \pi\rangle)$ and $\otimes_{i=1}^{k(n)} \frac{1}{\sqrt{2}}(|\sigma_i\rangle - |\sigma_i \circ \pi\rangle)$, where $\sigma_i \in \mathbb{S}_n, \pi \in \mathbb{K}_n$ are permutations ($\mathbb{K}_n$ is the set of all permutation in $\mathbb{S}_n$ of order 2, such that $\pi(i) \neq i$ for all i), and k is some polynomial, such that for each state the array of σ_i's is chosen at random. In the same paper, the authors discuss how one can explore the indistinguishability of these quantum states in the scope of cryptography and propose a cryptographic public key scheme that, under the hardness assumption of QSCD_{ff}, is secure against polynomial quantum adversaries using π as a trapdoor.

The rest of this paper is organized as follows: In the next section we present the basic notions, problems, notation, and results used in the rest of the paper. We also show the equivalence between a previously known algorithm, used in [41], and a particular orthogonal measurement that is used to prove the security of our protocol, under the assumption that QSCD_{ff} is polynomially hard for quantum computers. In Sec. III we present our protocol for a bit-string oblivious transfer and analyze its correctness and security. In Sec. IV we present the conclusions.

II. PRELIMINARIES

We use a binary alphabet $\Sigma = \{0,1\}$ and strings of length ℓ, which are elements of Σ^ℓ. The group $\mathbb{S}_n$ is the set of all permutations over the set $\{1,\ldots,n\}$, whose elements we denote by Greek letters $\sigma,\tau,\delta,\ldots$, together with the composition operation $\circ$ [42]. In the rest of the paper we assume n to be of the form $2(2m+1)$, for some $m \in \mathbb{N}$.

Example 1. Consider $n = 6$ and $\sigma \in \mathbb{S}_6$ defined as $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 1, \sigma(4) = 5, \sigma(5) = 4$, and $\sigma(6) = 6$. We represent this permutation as $\sigma = (1\ 2\ 3)(4\ 5)$, where $(1\ 2\ 3)$ and $(4\ 5)$ represent the orbits of elements of $\{1,\ldots 6\}$. The orbit of $i \in \{1,\ldots,n\}$ with respect to σ is $(i\ \sigma(i)\ \ldots\ \sigma^j(i))$, where the superscript is the number of times σ is applied to element i, and j is the smallest integer such that $\sigma(\sigma^j(i)) = i$.

Notice that this representation is not unique, and we can represent this same σ as $(4\ 5)(1\ 2\ 3)$ or $(5\ 4)(2\ 3\ 1)$.

Given a permutation $\sigma \in \mathbb{S}_n$, other than the identity, one can decompose it into a sequence of transpositions, i.e., elementary permutations that only exchange two elements. It is easy to see that such decomposition is not unique, but the number of transpositions, denoted by $\#(\sigma)$, has always the same *parity* and hence one can define the sign of a permutation σ as $\mathrm{sgn}(\sigma) = (-1)^{\#(\sigma)}$.

Example 2. Consider σ as defined in Example 1. Three possible decompositions of σ in terms of transpositions (derived from the three given representations) are $(1\ 3)(1\ 2)(4\ 5)$

and (4 5)(1 3)(1 2) and (5 4)(2 1)(2 3), and all of them have parity 1.

One can, in fact, show that if a permutation σ generates L orbits of elements from $\{1,\ldots,n\}$ of lengths $\ell_1,\ldots,\ell_L$ then $\#(\sigma)=\sum_{i=1}^{L}(\ell_i-1)$. In our case, the representation of σ in Example 1 has orbits of length 3 and 2, hence $\#(\sigma)=(3-1)+(2-1)=3$. The same σ was represented in this example with three orbits of length 2, hence $\#(\sigma)=3\times(2-1)=3$.

Since $|\mathbb{S}_n|=n!$ one needs $\log(n!)=\sum_{i=1}^{n}\log i\leqslant n\log n\in O(n\log n)$ bits to represent each $\sigma\in\mathbb{S}_n$. Note that $\mathbb{S}_n$ consists of two sets of equal size: $\mathbb{E}_n$ containing the even permutations (i.e., permutations with sign 1), and its complement $\mathbb{O}_n$ consisting of all odd permutations. Hence $\mathbb{S}_n=\mathbb{E}_n\cup\mathbb{O}_n$.

As in Ref. [41], we consider the following subset of $\mathbb{S}_n$:

$$\mathbb{K}_n=\{\pi\in\mathbb{S}_n:\pi\circ\pi=id_n \text{ and } \pi(i)\neq i,\forall i\in\{1,\ldots,n\}\}.$$

Example 3. One can see that the permutation σ of Example 1 is not in $\mathbb{K}_6$ as $\sigma\circ\sigma\neq id_n$, *e.g.*, $\sigma\circ\sigma(1)=\sigma(2)=3\neq 1$. We also do not have $\sigma(i)\neq i$ for $i=6$.

As an example of a permutation in $\mathbb{K}_6$ we have $\pi=(1\ 2)(3\ 6)(4\ 6)$.

One can immediately see that for an even n, any permutation $\pi\in\mathbb{K}_n$ can always be decomposed into $n/2$ transpositions since each element $i\in\{1,\ldots,n\}$ has to appear in (at least) one transposition [otherwise $\pi(i)=i$], and appears only once, as $\pi^2=id_n$. Given that we assumed $n=2(2m+1)$, we have that $\pi\in\mathbb{K}_n$ if and only if it can be written as an odd number of transpositions and hence $\pi\in\mathbb{O}_n$.

A counting argument revels that the size of $\mathbb{K}_n$ is

$$|\mathbb{K}_n|=\binom{n}{n/2}\left(\frac{n}{2}\right)!=\frac{n!}{\left(\frac{n}{2}\right)!}.$$

Let $n\in\mathbb{N}$ and $\pi\in\mathbb{K}_n$, and consider the Hilbert space $\mathcal{H}_n=\text{span}\{|\sigma\rangle:\sigma\in\mathbb{S}_n\}$, such that for all σ and σ', $\langle\sigma|\sigma'\rangle=\delta_{\sigma,\sigma'}$. Let $|\psi_\pi^{\pm}(\sigma)\rangle=\frac{1}{\sqrt{2}}(|\sigma\rangle\pm|\sigma\circ\pi\rangle)$. Notice that for every π the set $\{|\psi_\pi^{\pm}(\sigma)\rangle:\sigma\in\mathbb{E}_n\}$ is an orthonormal basis.

Indeed, we have that $\langle\psi_\pi^{\pm}(\sigma)|\psi_\pi^{\pm}(\sigma')\rangle=\frac{1}{2}(\langle\sigma|\sigma'\rangle\pm\langle\sigma|\sigma'\circ\pi\rangle\pm\langle\sigma\circ\pi|\sigma'\rangle+\langle\sigma\circ\pi|\sigma'\circ\pi\rangle)$ is either 1, when $\sigma=\sigma'$, or 0 otherwise. This is because $\sigma,\sigma'\in\mathbb{E}_n$ and $\pi\in\mathbb{O}_n$, and consequently $\langle\sigma|\sigma'\circ\pi\rangle=\langle\sigma\circ\pi|\sigma'\rangle=0$.

On the other hand $\langle\psi_\pi^{\pm}(\sigma)|\psi_\pi^{\mp}(\sigma')\rangle=\frac{1}{2}(\langle\sigma|\sigma'\rangle\mp\langle\sigma|\sigma'\circ\pi\rangle\pm\langle\sigma\circ\pi|\sigma'\rangle-\langle\sigma\circ\pi|\sigma'\circ\pi\rangle)$ is always 0: When $\sigma=\sigma'$, the first and last terms cancel each other, while the other two terms are always zero. Since $|\psi_\pi^{\pm}(\sigma)\rangle=\pm|\psi_\pi^{\pm}(\sigma\circ\pi)\rangle$, the set $\{|\psi_\pi^{\pm}(\sigma)\rangle:\sigma\in\mathbb{O}_n\}$ is also an orthonormal basis.

Consider the following quantum states defined in Ref. [41]

$$\begin{aligned}\rho_\pi^{+}&=\frac{1}{2n!}\sum_{\sigma\in\mathbb{S}_n}(|\sigma\rangle+|\sigma\circ\pi\rangle)(\langle\sigma|+\langle\sigma\circ\pi|)\\&=\frac{1}{n!}\sum_{\sigma\in\mathbb{S}_n}|\psi_\pi^{+}(\sigma)\rangle\langle\psi_\pi^{+}(\sigma)|\\&=\frac{2}{n!}\sum_{\sigma\in\mathbb{E}_n}|\psi_\pi^{+}(\sigma)\rangle\langle\psi_\pi^{+}(\sigma)|,\end{aligned}$$

and

$$\begin{aligned}\rho_\pi^{-}&=\frac{1}{2n!}\sum_{\sigma\in\mathbb{S}_n}(|\sigma\rangle-|\sigma\circ\pi\rangle)(\langle\sigma|-\langle\sigma\circ\pi|)\\&=\frac{1}{n!}\sum_{\sigma\in\mathbb{S}_n}|\psi_\pi^{-}(\sigma)\rangle\langle\psi_\pi^{-}(\sigma)|\\&=\frac{2}{n!}\sum_{\sigma\in\mathbb{E}_n}|\psi_\pi^{-}(\sigma)\rangle\langle\psi_\pi^{-}(\sigma)|.\end{aligned}$$

We are interested in these particular states as they are orthogonal to each other and hence fully distinguishable, *provided one knows* which π was used to prepare them. Without the knowledge of π the problem of distinguishing these states is believed to be polynomially hard even for quantum computers, as stated in Ref. [41]. First, we state the problem.

Problem 1 (QSCD_{ff}*).* The Quantum State Computational Distinction with Fully Flipped Permutations Problem, denoted by QSCD_{ff}, is defined as

Instances. Two quantum states $(\rho_\pi^{+})^{\otimes k(n)}$ and $(\rho_\pi^{-})^{\otimes k(n)}$ where $n=2(2m+1)$ for some $m\in\mathbb{N}$, and k is some fixed polynomial, *i.e.*, each state consists of $k(n)$ copies of ρ_π^{+} and ρ_π^{-}, respectively.

Question. Are $(\rho_\pi^{+})^{\otimes k(n)}$ and $(\rho_\pi^{-})^{\otimes k(n)}$ computationally indistinguishable if π is unknown, *i.e.*, is the probability to distinguish between the states $(\rho_\pi^{+})^{\otimes k(n)}$ and $(\rho_\pi^{-})^{\otimes k(n)}$ a negligible function, when running a quantum polynomial time algorithm without knowing π [43]?

The problem QSCD_{ff} is closely related to the hidden subgroup problem over symmetric groups for which no one knows an efficient quantum algorithm to solve it. In Ref. [41] the authors reduced the QSCD_{ff} problem to a variant of the unique graph automorphism problem, that is also presumably hard for quantum computers with polynomial time resources, proving the following hardness result.

Theorem 1 ([41]). If there exists a polynomial-time quantum algorithm that solves QSCD_{ff} with non-negligible advantage, then there exists a polynomial-time quantum algorithm that solves the graph automorphism problem in the worst case for infinitely many input lengths.

The interesting property that makes this problem suitable for cryptography is that it has a trapdoor that allows one to efficiently distinguish the states: One can distinguish, with certainty, $(\rho_\pi^{+})^{\otimes k(n)}$ from $(\rho_\pi^{-})^{\otimes k(n)}$, provided that an extra piece of information is given, in this case π. Furthermore, when using a permutation π' different from the trapdoor π, the probability of distinguishing these states is the same as plain guessing. In order to present the protocol and analyze its complexity and security, we need the following obvious proposition:

Proposition 1. The following linear operators are unitary (for all $\sigma,\tau,\pi\in\mathbb{S}_n$):

$$C_{\circ}^{l}=\sum_{\sigma,\tau}|\sigma,\sigma\circ\tau\rangle\langle\sigma,\tau|;$$

$$C_{\circ}^{r}=\sum_{\sigma,\tau}|\sigma,\tau\circ\sigma\rangle\langle\sigma,\tau|;$$

$$C_\pi = |0\rangle\langle 0| \otimes \sum_\sigma |\sigma\rangle\langle\sigma| + |1\rangle\langle 1| \otimes \sum_\sigma |\sigma \circ \pi\rangle\langle\sigma|;$$

$$C_1 = \mathbb{1}_2 \otimes \sum_{\sigma \neq \pi} |\sigma\rangle\langle\sigma| + (|1\rangle\langle 0| + |0\rangle\langle 1|) \otimes |\pi\rangle\langle\pi|;$$

$$C_{\text{swap}} = \sum_{\sigma,\tau} |\tau,\sigma\rangle\langle\sigma,\tau|;$$

$$C_{\text{sgn}} = \sum_\sigma (-1)^{\text{sgn}(\sigma)} |\sigma\rangle\langle\sigma|.$$

It is easy to see by explicit calculation that each operator C from the above proposition satisfies the unitarity condition $CC^\dagger = C^\dagger C = \mathbb{1}$ (here, as well as in the rest of the paper, $\mathbb{1}$ is the identity operator in the corresponding space).

The first two operations, $C_\circ^l$ and $C_\circ^r$, leave the state of the first register unchanged, while the state of the second is transformed into the state given by the "left" and the "right" composition of the two permutations, respectively. Controlled operation C_π composes the second-register permutation with π from the right, while C_1 inverts the basis states $|0\rangle$ and $|1\rangle$ of the first register if the second is in the state $|\pi\rangle$ (both operations are defined for an arbitrary permutation, but in the paper we will constrain to cases $\pi \in \mathbb{K}_n$). Finally, C_{swap} swaps the states of the two registers, while C_{sgn} changes the sign in front of the basis vectors $|\sigma\rangle$.

The quantum algorithm presented in Ref. [41], Algorithm 1, justifies that given the private key π one can produce the states $(\rho_\pi^+)^{\otimes k(n)}$ in *polynomial time* using only Hadamard H and the operations defined in Proposition 1.

Algorithm 1 (to generate $(\rho_\pi^+)^{\otimes k(n)}$ [41]). For the sake of simplicity of the presentation we consider a fixed n and the case where $k(n) = 1$. The reader can easily generalize the argument for $k(n)$ systems.

Input: $\pi \in \mathbb{K}_n$.

Output: $\rho_\pi^+ = \dfrac{2}{n!} \displaystyle\sum_{\sigma \in \mathbb{E}_n} (|\sigma\rangle + |\sigma \circ \pi\rangle)(\langle\sigma| + \langle\sigma \circ \pi|)$.

Step 1. *Select a random $\sigma \in \mathbb{S}_n$ and prepare the initial state* $|0\rangle \otimes |\text{id}_n\rangle \otimes |\sigma\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^{n \log n} \otimes \mathbb{C}^{n \log n}$.
Step 2. *Apply* $H \otimes \mathbb{1} \otimes \mathbb{1}$.
Step 3. *Apply* $C_\pi \otimes \mathbb{1}$.
Step 4. *Apply* $C_1 \otimes \mathbb{1}$.
Step 5. *Apply* $\mathbb{1} \otimes C_{\text{swap}}$.
Step 6. *Apply* $\mathbb{1} \otimes C_\circ^r$.

The third register is now in the state $|\psi_\pi^+(\sigma)\rangle = \frac{1}{\sqrt{2}}(|\sigma\rangle + |\sigma \circ \pi\rangle)$. Since σ is chosen at random, the ensemble of such systems is represented by the mixed state ρ_π^+.

Note that it is possible to prepare the state ρ_π^+ as a partial trace of an entangled state. In order to do that, one should start from $|0\rangle|\text{id}_n\rangle \frac{1}{\sqrt{n!}} \sum_{\sigma \in \mathbb{S}_n} |\sigma\rangle$. Upon applying Steps 1–6 of the previous algorithm the overall state is $|0\rangle \frac{1}{\sqrt{n!}} \sum_{\sigma \in \mathbb{S}_n} |\sigma\rangle \frac{1}{\sqrt{2}}(|\sigma\rangle + |\sigma \circ \pi\rangle)$ and the third register is again in the state ρ_π^+.

In the next lemma we show that one can easily transform ρ_π^+ into ρ_π^-, and vice versa, without knowing π. This property follows immediately from the fact that π is an odd permutation.

Lemma 1 ([41]). There exists a *polynomial-time* quantum algorithm (in fact C_{sgn}) that, with probability 1, transforms ρ_π^+ into ρ_π^- and keeps completely mixed state $\rho_{\mathbb{1}} = \frac{1}{n!} \sum_{\sigma \in \mathbb{S}_n} |\sigma\rangle\langle\sigma|$ invariant, for any $n = 2(2m+1)$, with $m \in \mathbb{N}$, and any permutation $\pi \in \mathbb{K}_n$.

Indeed, for a pure $|\psi_\pi^+(\sigma)\rangle$, with $\pi \in \mathbb{K}_n$, by applying C_{sgn} we get

$$\begin{aligned} |\psi_\pi^+(\sigma)\rangle &= \frac{|\sigma\rangle + |\sigma \circ \pi\rangle}{\sqrt{2}} \\ &\xrightarrow{C_{\text{sgn}}} \frac{(-1)^{\text{sgn}(\sigma)}|\sigma\rangle + (-1)^{\text{sgn}(\sigma)+1}|\sigma \circ \pi\rangle}{\sqrt{2}} \\ &= (-1)^{\text{sgn}(\sigma)} |\psi_\pi^-(\sigma)\rangle. \end{aligned}$$

Notice that determining the sign of a permutation is a computation that can be done in polynomial time. Furthermore, it is easy to see that this algorithm leaves $\rho_{\mathbb{1}}$ invariant.

The following algorithm explores the trapdoor property of QSCD$_{ff}$: Given π, one can distinguish in *polynomial time* and with probability 1 the quantum states $(\rho_\pi^+)^{\otimes k(n)}$ and $(\rho_\pi^-)^{\otimes k(n)}$ (see [41] for the details of the proof).

Algorithm 2 (distinguishing $(\rho_\pi^+)^{\otimes k(n)}$ from $(\rho_\pi^-)^{\otimes k(n)}$). For the sake of simplicity of the presentation we consider a fixed n and the case where $k(n) = 1$. The reader can easily generalize the argument for $k(n)$ systems.

Input: $\pi \in \mathbb{K}_n$ *and a quantum state* χ_π *that is either* ρ_π^+ *or* ρ_π^-.
Output: 0 if $\chi_\pi = \rho_\pi^+$, and 1 if $\chi_\pi = \rho_\pi^-$.
Step 1. *Prepare the system in the state* $|0\rangle\langle 0| \otimes \chi_\pi$, *where* $|0\rangle \in \mathbb{C}^2$.
Step 2. *Apply* $H \otimes \mathbb{1}$.
Step 3. *Apply* C_π.
Step 4. *Apply again* $H \otimes \mathbb{1}$.
Step 5. *Measure* $M_+ = (0 \cdot |0\rangle\langle 0| + 1 \cdot |1\rangle\langle 1|) \otimes \mathbb{1}$ *and output the result.*

After performing the second step the obtained state is $\frac{1}{2}|+\rangle\langle+| \otimes \chi_\pi$, with $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Since $C_\pi(|+\rangle(|\sigma\rangle \pm |\sigma \circ \pi\rangle)) = |\pm\rangle(|\sigma\rangle \pm |\sigma \circ \pi\rangle)$, the overall state after the third step is $|\pm\rangle\langle\pm| \otimes \rho_\pi^\pm$. So, by applying Hadamard on the first register and measuring it in the computational basis we get the desired outcome.

We have shown that the above algorithm is able to distinguish with certainty between $\chi_\pi = (\rho_\pi^+)^{\otimes k(n)}$ and $\chi_\pi = (\rho_\pi^-)^{\otimes k(n)}$, using only polynomial quantum resources, provided that we use the correct π. We will show later that if $\chi_{\pi'}$ is created with a different permutation $\pi' \neq \pi$, then the answer will be a random variable with distribution close to uniform (see Theorem 3 for details).

Algorithm 2 provides a computational approach to the problem of distinguishing the two states and is suitable for defining the public key scheme presented in Ref. [41]. In our work, due to the nature of the problem at hand, we will instead consider measurements. We will show next that the result of Algorithm 2 is equivalent to measuring the orthogonal observable,

$$M_\pi = 0 \cdot P_\pi^+ + 1 \cdot P_\pi^-,$$

where

$$P_\pi^\pm = \sum_{\sigma\in\mathbb{E}_n} |\psi_\pi^\pm(\sigma)\rangle\langle\psi_\pi^\pm(\sigma)| = \frac{1}{2}\sum_{\sigma\in\mathbb{E}_n}(|\sigma\rangle \pm |\sigma\circ\pi\rangle)(\langle\sigma| \pm \langle\sigma\circ\pi|).$$

Proposition 2. Applying Algorithm 2 with π and measuring $M_\pi = 0\cdot P_\pi^+ + 1\cdot P_\pi^-$ are equivalent processes.

The probability distribution of the outcomes of the Algorithm 2 is the same as the probability distribution of the outcomes of the measurement M_π, and the resulting states are the same.

As in Algorithm 2, we consider a fixed n and the case where $k(n) = 1$. The reader can easily generalize the argument for $k(n)$ systems using $M_\pi^{\otimes k(n)}$ instead of M_π.

Proof. In order to prove this proposition, first notice that by the linearity of the measurement, it is enough to show the result only for the case of a general pure $|\psi\rangle$. Since for every π the set $\{|\psi_\pi^\pm(\sigma)\rangle : \sigma\in\mathbb{E}_n\}$ is a basis, we can write

$$|\psi\rangle = \sum_{\sigma\in\mathbb{E}_n}(c_\sigma^+|\psi_\pi^+(\sigma)\rangle + c_\sigma^-|\psi_\pi^-(\sigma)\rangle).$$

We first compute the following:

$$P_\pi^\pm|\psi\rangle = \sum_{\sigma\in\mathbb{E}_n}|\psi_\pi^\pm(\sigma)\rangle\langle\psi_\pi^\pm(\sigma)|\left(\sum_{\sigma'\in\mathbb{E}_n}(c_{\sigma'}^+|\psi_\pi^+(\sigma')\rangle + c_{\sigma'}^-|\psi_\pi^-(\sigma')\rangle)\right)$$
$$= \sum_{\sigma,\sigma'\in\mathbb{E}_n}|\psi_\pi^\pm(\sigma)\rangle(c_{\sigma'}^+\langle\psi_\pi^\pm(\sigma)|\psi_\pi^+(\sigma')\rangle + c_{\sigma'}^-\langle\psi_\pi^\pm(\sigma)|\psi_\pi^-(\sigma')\rangle) = \sum_{\sigma\in\mathbb{E}_n}c_\sigma^\pm|\psi_\pi^\pm(\sigma)\rangle.$$

The result of the first four steps of Algorithm 2 is ($|\psi_\pi^\pm(\sigma)\rangle = \pm|\psi_\pi^\pm(\sigma\pi)\rangle$):

$$(H\otimes\mathbb{1})C_\pi(H\otimes\mathbb{1})|0\rangle|\psi\rangle = \frac{1}{\sqrt{2}}(H\otimes\mathbb{1})C_\pi(|0\rangle+|1\rangle)|\psi\rangle$$
$$= \frac{(H\otimes\mathbb{1})}{\sqrt{2}}C_\pi\left(|0\rangle|\psi\rangle + |1\rangle\sum_{\sigma\in\mathbb{E}_n}(c_\sigma^+|\psi_\pi^+(\sigma)\rangle + c_\sigma^-|\psi_\pi^-(\sigma)\rangle)\right)$$
$$= \frac{(H\otimes\mathbb{1})}{\sqrt{2}}\left(|0\rangle|\psi\rangle + |1\rangle\sum_{\sigma\in\mathbb{E}_n}(c_\sigma^+|\psi_\pi^+(\sigma)\rangle - c_\sigma^-|\psi_\pi^-(\sigma)\rangle)\right)$$
$$= \frac{|0\rangle+|1\rangle}{2}\sum_{\sigma\in\mathbb{E}_n}(c_\sigma^+|\psi_\pi^+(\sigma)\rangle + c_\sigma^-|\psi_\pi^-(\sigma)\rangle) + \frac{|0\rangle-|1\rangle}{2}\sum_{\sigma\in\mathbb{E}_n}(c_\sigma^+|\psi_\pi^+(\sigma)\rangle - c_\sigma^-|\psi_\pi^-(\sigma)\rangle)$$
$$= |0\rangle\sum_{\sigma\in\mathbb{E}_n}c_\sigma^+|\psi_\pi^+(\sigma)\rangle + |1\rangle\sum_{\sigma\in\mathbb{E}_n}c_\sigma^-|\psi_\pi^-(\sigma)\rangle = |0\rangle P_\pi^+|\psi\rangle + |1\rangle P_\pi^-|\psi\rangle.$$

Measuring the first register in the computational basis (Step 5 of Algorithm 2) collapses the second register (up to a normalization factor) in the state $P_\pi^\pm|\psi\rangle$ with probability $||P_\pi^\pm|\psi\rangle||^2$, just as if M_π was measured. ■

An immediate corollary is that M_π distinguishes ρ_π^+ and ρ_π^- with probability one.

At the end of an oblivious transfer protocol, Bob must know if he received the message or not. In our protocol, this step is guaranteed by a *universal hash function*. These functions map larger strings to strings of smaller size, hence collisions are unavoidable, i.e., different messages can be mapped to the same hash. Despite this fact, one can design these functions in such a way that

(1) they are computationally efficient, i.e., one can compute its value from a message in polynomial time;

(2) hashes are almost equally distributed.

Definition 1. Let A and B be two sets of size a and b, respectively, such that $a > b$, and let $\mathbb{H}$ be a collection of hash functions $h : A \to B$. $\mathcal{H}$ is said to be a *universal family of hash functions* if, for all $x, y \in A$,

$$\Pr_{h\in\mathbb{H}}[h(x) = h(y)] \leqslant \frac{1}{b}.$$

A straightforward consequence of this definition is the following theorem.

Theorem 2. Let A and B be two sets of size a and b, respectively, such that $a > b$ and let $\mathbb{H}$ be a collection of hash functions $h : A \to B$.

If $\mathbb{H}$ is a universal family of hash functions then for any set $A' \subset A$ of size N and for any $x \in A$, the expected number of collisions between x and other elements in A' is at most N/b.

Notice that, in particular, if we request A to contain all strings of length ℓ and B to have length say $\omega = \lfloor\sqrt{\ell}\rfloor$, then the number of expected collisions is $2^\omega = 2^{\lfloor\sqrt{\ell}\rfloor}$, hence the probability of finding a collision is negligible in ℓ. There are several standard ways to construct universal families of hash functions (see [44] for examples and details). In this paper, we prove the security of our protocol for a specific universal family

of hash functions that are easy to construct. For simplicity, we postpone the description of this rather technical detail for the proof of Theorem 4.

III. THE BIT-STRING OBLIVIOUS TRANSFER PROTOCOL

In this section, we present a quantum protocol that achieves oblivious transfer of a bit string from Alice to Bob in polynomial time. As already mentioned in the introduction, oblivious transfer protocol is a protocol in which Alice sends (transfers) a message $m = m_1 m_2 \dots m_\ell$ of length ℓ to Bob (during the transfer phase), which is recovered by him with only 50% of probability during the opening phase (the *transfer is probabilistic*). The protocol must satisfy two additional properties: be *concealing*, i.e., Bob must not learn m before the opening phase; and be *oblivious*, i.e., after the opening phase Alice must *not know with certainty* if Bob received m or not. Notice that Bob, unlike Alice, *does know* at the end of the protocol if he received the intended message. In the rest of this section, we present our results in terms of orthogonal measurements M_π, which according to Proposition 2 can be realized in polynomial time by applying Algorithm 2.

Protocol 1 (Bit-string oblivious transfer)

Security parameters ℓ and $n \geqslant \ell$.
Message to transfer $m = m_1 \dots m_\ell$.
Universal hash function $h : \Sigma^\ell \to \Sigma^\omega$.
Secret key $\pi \in \mathbb{K}_n$.
Transferring phase:
Step 1. Alice generates uniformly at random the secret key $\pi \in \mathbb{K}_n$.
Step 2. Using Algorithm 1 with π, and the operation C_{sgn}, she constructs the state $\rho_\pi^m = \rho_\pi^{m_1} \otimes \rho_\pi^{m_2} \otimes \dots \otimes \rho_\pi^{m_\ell}$ where $\rho_\pi^{m_i} = \rho_\pi^+$ if $m_i = 1$, and $\rho_\pi^{m_i} = \rho_\pi^-$ otherwise.
Step 3. Alice sends ρ_π^m and $y = h(m)$ to Bob.
Opening phase:
Step 4. Bob generates uniformly at random $\tau \in \mathbb{S}_n$ and sends it to Alice.
Step 5. Alice computes uniformly at random either $\delta = \pi \circ \tau$ or $\delta = \tau \circ \pi$, and sends it back to Bob.
Step 6. Bob computes uniformly at random either $\gamma = \delta \circ \tau^{-1}$ or $\gamma = \tau^{-1} \circ \delta$.
Step 7. Bob measures the observable $M_\gamma^{\otimes \ell}$ on the system given by Alice, obtaining the result $\tilde{m}$.
Step 8. Bob checks if $h(\tilde{m}) = y$. If so, he concludes that the message sent by Alice is $\tilde{m}$, *i.e.*, $\tilde{m} = m$.
Step 9. If Bob got the correct message, he chooses another $\pi' \in \mathbb{K}_n$ and measures the observable $M_{\pi'}^{\otimes \ell}$ on the system, obtaining result $r = r_1 \dots r_\ell$. If approximately half of the results r_i are different from the corresponding m_i, then Bob accepts the message; otherwise he aborts the protocol declaring that Alice tried to cheat.

Below, we present an informal analysis of the security of the protocol, while the rigorous security criterion is given in the statement of Theorem 4. As discussed in the Introduction, in this paper we only consider the security of the protocol *per se*, while the security of complex schemes that use our protocol as a black box subroutine is to be carried out when analyzing particular realizations of such schemes. The protocol's security depends on parameter n through Theorem 1, which is satisfied asymptotically, i.e., for *sufficiently large* values of n. The dependence on ℓ comes through probabilistic transfer requirement (see below Point 2 of Theorems 3 and 4), stated in terms of negligible functions $\varepsilon(\ell)$, again for *sufficiently large* values of ℓ. [A function $\varepsilon(\ell)$ is said to be negligible if $\varepsilon(\ell) < 1/p(\ell)$ for any polynomial $p(\ell)$ and sufficiently large ℓ.] Note that, since we require that $n \geqslant \ell$ [45], the security of the protocol is stated in terms of a negligible function in ℓ: Any security bound proven for n is at the same time a bound for ℓ (recall that a function is negligible if it scales to zero faster than the inverse of any polynomial, *for sufficiently large* value of its argument).

To show that our proposal is a bit-string oblivious transfer protocol we must prove the following:

(1) If Alice is honest, then the protocol is computationally *concealing*, i.e., Bob cannot learn the message m before the opening phase. Notice that this follows directly from the hardness assumption of computational indistinguishability of states ρ^+ and ρ^- (Theorem 1) and from the fact that the hash function used for comparison has exponentially many collisions, i.e., only with negligible probability Bob can correctly invert h in probabilistic polynomial time and obtain m.

(2) If Alice is honest, then roughly in 50% of the cases Bob will obtain m (*probabilistic transfer*). Notice that the acknowledgment that $\tilde{m}$ is correct is given by the comparison of y with $h(\tilde{m})$.

(3) If Bob is honest, then the protocol is *oblivious*, i.e., Alice cannot learn with certainty whether Bob got the message m or not. This is a consequence of the impossibility of faster than light information transmission and it is ensured by the last step of the protocol. The *rationale* for Step 9 is to prevent Alice from cheating by sending a state that would allow her to know with certainty that Bob would receive the message. We postpone the discussion of this point for the end of the paper.

Notice that if the order of compositions agree in Steps 5 and 6 of the protocol, i.e., Alice and Bob applied, respectively, τ and τ^{-1} on the same side, then $\gamma = \pi$; otherwise, if Alice and Bob applied τ and τ^{-1} on different sides, then both $\gamma = \tau^{-1} \circ \pi \circ \tau$ and $\gamma = \tau \circ \pi \circ \tau^{-1}$ belong to $\mathbb{K}_n$ but are different from π (in fact, $\tau^{-1} \circ \pi \circ \tau = \tau \circ \pi \circ \tau^{-1}$).

Also, due to the properties of hash functions, the probability of two randomly chosen messages having the same hash value is negligible. Therefore, if in Step 8 of the protocol $h(\tilde{m}) = y = h(m)$, then $\tilde{m}$ is, up to negligible probability, the actual message m.

Therefore, if both Alice and Bob are honest, then Bob measures with equal probability either $M_\pi^{\otimes \ell}$ or $M_{\pi'}^{\otimes \ell}$, where $\pi' = \tau \circ \pi \circ \tau^{-1} \in \mathbb{K}_n$ for some $\tau \in \mathbb{S}_n$. If he measures $M_\pi^{\otimes \ell}$, then the measurement result $\tilde{m}$ will indeed be equal to m, which he can confirm by comparing $h(\tilde{m})$ and y. Otherwise, and since π' does not match the used π, each single-system measurement will yield a random $\tilde{m}_i$, and by confirming that $h(\tilde{m}) \neq y$, Bob will know that he did not receive the intended message.

This way, if both parties are honest, Bob will recover the message with probability 1/2. We formalize these in the following theorem.

Theorem 3 (Correctness of Protocol 1). Assume that QSCD_{ff} is polynomially hard even for quantum computers. If Alice and Bob correctly run Protocol 1 to transfer message $m = m_1 \dots m_\ell$ from Alice to Bob, then

(1) *(Concealing)* Bob cannot infer m before the opening phase except with negligible probability on ℓ, the size of the message m.

(2) *(Probabilistic transfer)* Bob will receive m with probability $1/2 + \varepsilon(\ell)$, where $\varepsilon(\ell)$ is a negligible function.

(3) *(Oblivious)* Alice remains oblivious to the fact that Bob received the message.

Proof. We prove each item stated in the theorem separately:

(1) Since Bob is honest, he does not perform any quantum operations other than those prescribed by the protocol. In particular, he does not perform any measurement on the quantum system received by Alice before the opening phase. On the other hand, since it is assumed that h is a universal hash function, trying to recover m by computing the inverse of h is only possible with negligible probability. In fact, since the hash function h maps strings of length ℓ to strings of length $\omega = \lfloor\sqrt{\ell}\rfloor$, by Theorem 2 one can recover m with probability $2^{-\lfloor\sqrt{\ell}\rfloor}$ which is negligible in ℓ (recall that by negligible we mean that the probability to recover m scales to zero faster than the inverse of any polynomial in the length of the message).

(2) It is easy to see that after Step 6 of Protocol 1, $\gamma = \pi$ with probability $1/2$ if both Alice and Bob make their choices in Steps 5 and 6 at random, and Bob chooses τ uniformly at random.

For $\gamma = \pi$, the result that the measurement $M_\pi^{\otimes\ell}(\rho_\pi^m)$ equals m follows from Proposition 2.

It remains to show that when $\gamma = \pi' \neq \pi$, by performing the measurement $M_\gamma^{\otimes\ell}$, the probability of recovering m from the quantum state ρ_π^m sent by Alice is negligible. First we prove that for any σ the result of the measurement of $M_{\pi'}^\ell$ is random. Observe that

$$
\begin{aligned}
P_{\pi'}^{\pm}(|\psi_\pi^+(\sigma)\rangle) &= \sum_{\sigma'\in\mathbb{E}_n} |\psi_{\pi'}^{\pm}(\sigma')\rangle\langle\psi_{\pi'}^{\pm}(\sigma')|\psi_\pi^+(\sigma)\rangle \\
&= \sum_{\sigma'\in\mathbb{E}_n} \frac{(|\sigma'\rangle \pm |\sigma'\circ\pi'\rangle)(\langle\sigma'| \pm \langle\sigma'\circ\pi'|)(|\sigma\rangle + |\sigma\circ\pi\rangle)}{2\sqrt{2}} \\
&= \sum_{\sigma'\in\mathbb{E}_n} \frac{(|\sigma'\rangle \pm |\sigma'\circ\pi'\rangle)(\langle\sigma'|\sigma\rangle + \langle\sigma'|\sigma\circ\pi\rangle \pm \langle\sigma'\circ\pi'|\sigma\rangle \pm \langle\sigma'\circ\pi'|\sigma\circ\pi\rangle)}{2\sqrt{2}} \\
&= \begin{cases} \displaystyle\sum_{\sigma'\in\mathbb{E}_n} \frac{(|\sigma'\rangle \pm |\sigma'\circ\pi'\rangle)(\langle\sigma'|\sigma\rangle + \langle\sigma'\circ\pi'|\sigma\circ\pi\rangle)}{2\sqrt{2}}, & \text{if } \sigma \in \mathbb{E}_n \\ \displaystyle\sum_{\sigma'\in\mathbb{E}_n} \frac{(|\sigma'\rangle \pm |\sigma'\circ\pi'\rangle)(\langle\sigma'|\sigma\circ\pi\rangle \pm \langle\sigma'\circ\pi'|\sigma\rangle)}{2\sqrt{2}}, & \text{if } \sigma \in \mathbb{O}_n \end{cases} \\
&= \frac{1}{2\sqrt{2}}(|\sigma\rangle \pm |\sigma\circ\pi'\rangle + |\sigma\circ\pi\rangle \pm |\sigma\circ\pi\circ\pi'\rangle) \\
&= \frac{1}{2}(|\psi_{\pi'}^{\pm}(\sigma)\rangle \pm |\psi_{\pi'}^{\pm}(\sigma\circ\pi\circ\pi')\rangle \\
&= \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}(|\psi_{\pi'}^{\pm}(\sigma)\rangle \pm |\psi_{\pi'}^{\pm}(\sigma\circ\pi\circ\pi')\rangle)\right).
\end{aligned}
$$

Since $|\psi_{\pi'}^{\pm}(\sigma)\rangle$ and $|\psi_{\pi'}^{\pm}(\sigma\circ\pi\circ\pi')\rangle$ are orthogonal, $\frac{1}{\sqrt{2}}(|\psi_{\pi'}^{\pm}(\sigma)\rangle \pm |\psi_{\pi'}^{\pm}(\sigma\circ\pi\circ\pi')\rangle)$ is a unit vector and $\|P_{\pi'}^{\pm}(\psi_\pi^+)\|^2 = 1/2$. Hence, the probability of recovering $\pm$ from ρ_π^+ is

$$
\mathrm{Prob}(+; \mathrm{M}_{\pi'}, \rho_\pi^+) = \mathrm{Tr}[P_{\pi'}^+ \rho_\pi^+ P_{\pi'}^+] = \frac{2}{n!}\sum_{\sigma\in\mathbb{E}_n} \mathrm{Tr}[P_{\pi'}^+ |\psi_{\pi'}^+(\sigma)\rangle\langle\psi_{\pi'}^+(\sigma)| P_{\pi'}^+] = \frac{2}{n!}\sum_{\sigma\in\mathbb{E}_n} \|P_{\pi'}^+(\psi_\pi^+)\|^2 = \frac{1}{2},
$$

and similarly $\mathrm{Prob}(-; M_{\pi'}, \rho_\pi^+) = \frac{1}{2}$. *Mutatis mutandis* we also have $\mathrm{Prob}(\pm; M_{\pi'}, \rho_\pi^-) = \frac{1}{2}$. Hence, by measuring $M_{\pi'}^{\otimes\ell}$ on ρ_π^m the probability of recovering m is negligible in ℓ.

To conclude the proof of the second item we need to show that Bob aborts the protocol in Step 9 only with negligible probability.

Notice that to reach Step 9, where Bob aborts the protocol, he must have run successfully the verification in Step 7. This implies that Bob performed the measurement with the correct trapdoor π, hence the state stayed invariant, i.e., it is still ρ_π^m. If Bob chooses in Step 9 random $\pi' \neq \pi$, then the probability of recovering each bit of the message, as just seen above, is equal to $1/2$ and so, by a simple binomial argument, the probability of having a significant difference from half of the states is negligible on ℓ.

(3) Notice that Alice must send a permutation δ, such that both $\delta\circ\tau^{-1}$ and $\tau^{-1}\circ\delta$ are from $\mathbb{K}_n$. Bob's choice to compose δ with τ^{-1} on the left, or on the right, is random and unknown to Alice: after sending δ there is no more communication between Alice and Bob and there is no information transmission from Bob to Alice. Therefore, the choice of Bob's measurement observable $M_\gamma^{\otimes\ell}$ is also unknown to Alice as

well: Alice cannot know if Bob has obtained the message m, or not. ■

Finally we prove the security of the protocol. In the description of the protocol, only the encrypted message is sent as a state of a quantum system (ρ_π^m), while the hash value $h(m)$ is sent as a classical information. Nevertheless, even if joint *quantum* measurements on both systems carrying the encrypted message and the hash value were allowed, this would not give advantage to a cheating Bob.

Let $\mathcal{H}_m$ and $\mathcal{H}_h$ be the Hilbert spaces of the systems carrying the encrypted message and the hash value $h(m)$, respectively. In the following, we show that performing joint measurements over the entire Hilbert space $\mathcal{H}_{mh} = \mathcal{H}_m \otimes \mathcal{H}_h$ cannot achieve better cheating efficiency than an optimal strategy involving only measurements on subsystems given by $\mathcal{H}_m$ and $\mathcal{H}_h$.

After the transferring phase, Bob's description of the joint system received from Alice is given by the quantum state (recall that hash values are bit strings of length ω):

$$\rho_B = \frac{1}{|\mathbb{K}_n|}\sum_\pi \frac{1}{2^k}\sum_m \rho_\pi^m \otimes |h(m)\rangle\langle h(m)|$$

$$= \frac{1}{|\mathbb{K}_n|}\sum_\pi \frac{1}{2^\omega}\sum_h \left(\frac{1}{2^{k-\omega}}\sum_{m_h\in h^{-1}} \rho_\pi^{m_h}\right)\otimes |h\rangle\langle h|.$$

Note that states $|h\rangle \in \mathcal{H}_h$ are orthogonal to each other (they are fully distinguishable).

The aim of a cheating Bob is to discriminate among $|\mathbb{K}_n|\cdot 2^k$ states $\{\rho_\pi^m \otimes |h(m)\rangle\langle h(m)|\}$. The optimal probability of success of the joint-measurement discrimination strategy is the expected success probability of the same strategy when applied on the systems from subensembles, each having the same hash value h: $p^J_{\text{succ}} = \sum_h 2^{-\omega} p^J_{\text{succ}}(h)$, where $p^J_{\text{succ}}(h)$ is the success probability of the joint-measurement strategy for the subensemble given by a fixed hash value h. Each such subensemble consists of $|\mathbb{K}_n|\cdot 2^{k-\omega}$ uncorrelated product states $\{\rho_\pi^{m_h}\otimes|h\rangle\langle h| : m_h \in h^{-1}\}$, with the hash value h being the same for all elements of the subensemble. Thus, the success probability of the joint-measurement strategy for each subensemble is at most the optimal success probability of discriminating between the states $\{\rho_\pi^{m_h}\}$ having the same hash value h. Since states $|h\rangle$ are orthogonal to each other, by performing the subsystem measurements only (first measuring onto $\mathcal{H}_h$ in the $\{|h\rangle\}$ basis, and then performing an optimal discrimination measurement onto $\mathcal{H}_m$), Bob can achieve the optimal probability of discriminating among the states $\{\rho_\pi^m \otimes |h(m)\rangle\langle h(m)|\}$.

Therefore, one can assume that Bob is allowed to perform quantum measurements only on the system that carries the encrypted message. Below, we prove the protocol's security. We emphasize that the concealing and oblivious properties are satisfied against arbitrary cheating strategies of Alice and Bob, while the transfer is probabilistic as long as Bob is restricted to perform a few-qubit coherent measurements.

Theorem 4 (Security). Assume that QSCD_{ff} is polynomially hard even for a quantum computer. Protocol 1 is secure against cheating, i.e.,

(1) *(Concealing)* If Alice is honest, then even if he cheats, Bob cannot in polynomial time learn m before the opening phase, unless with negligible probability on ℓ.

(2) *(Probabilistic transfer)* If Alice is honest, and Bob is restricted to few-qubit coherent measurements, then he will receive m with probability $1/2 + \varepsilon(\ell)$, where $\varepsilon(\ell)$ is a negligible function.

(3) *(Oblivious)* If Bob is honest, then Alice cannot learn with certainty if he received the message even if she tries to cheat.

Proof.

(1) Notice that upon receiving a system from Alice the only thing Bob can do in order to learn its state (and hence the message m) is to simply "look" at it, i.e., perform a measurement in order to distinguish between ρ_π^+ and ρ_π^- (without knowing π). The concealing property follows directly from Theorem 1 proved in Ref. [41], which states that distinguishing ρ_π^+ from ρ_π^- without knowing π is polynomially hard even for a quantum computer. In other words, under the assumption that QSCD_{ff} is polynomially hard for a quantum computer, only with negligible probability one can, *without knowing* π, prepare in polynomial time a measurement that can distinguish ρ_π^+ from ρ_π^- with non-negligible advantage (i.e., the two states are *(quantum) computationally indistinguishable*). Moreover, if our Protocol 1 were not concealing, i.e., there would exist a (quantum) polynomial-time strategy that distinguishes ρ_π^+ from ρ_π^-, which could be used to devise an attack to the crypto-system developed in Ref. [41]. This is strongly believed to be false, since otherwise the power of quantum computers would be much stronger than it is conjectured by the community. For this reason, it is enough to show that Bob cannot learn (unless with negligible probability) either π, or the message m.

Notice that since Alice does not send π, the probability that Bob guesses the correct π is negligible on ℓ (since in the protocol we assume $n \geqslant \ell$): the size of $\mathbb{K}_n$ is

$$|\mathbb{K}_n| = \frac{n!}{(n/2)!},$$

which is already for $n = 10$ huge enough for practical purposes. Note that the additional information provided by Alice, the hash value $h(m)$, cannot help Bob recovering the message m, for a suitable chosen universal family of hash functions. First, we prove the concealing property using a particular hash function. Then, we show that the same is valid for a whole universal family of hash functions.

The message m can be divided into $\omega = \lfloor\sqrt{\ell}\rfloor$ consecutive blocks of bits $\bar{m}_i$ ($i = 1,\ldots,\omega$), each of the same length $\lfloor\sqrt{\ell}\rfloor$: $m = \bar{m}_1 \ldots \bar{m}_\omega$. The bits $h_i(m)$ of the hash function are given by the parity of the ith block of the message m: $h_1(m) = m_1 \oplus \ldots \oplus m_\omega$, etc. Note that each bit of $h(m)$ is independent of each other and depends only on specific block $\bar{m}_i$ of size ω. Therefore, if $h(m)$ would allow one to recover m with some non-negligible probability p, then $h_1(m)$ alone would help to recover $\bar{m}_1 = m_1 \oplus \ldots \oplus m_\omega$, with the same probability p. Such cheating strategy is impossible, assuming that the cryptographic system presented in Ref. [41] is secure.

Note that if a cryptographic system is computationally secure for coding a message m of length ℓ, then it must be computationally secure when encrypting a polynomially shorter message, say $\bar{m}_1$. Then, by guessing the parity of $\bar{m}_1$, and using the above cheating strategy, one would be able to break the crypto-system with non-negligible probability $p/2$.

From the above hash function h, it is easy to define the whole universal family of hash functions, indexed by permutations from $\mathbb{S}_\ell$, for which the concealing property is satisfied. Given $\alpha \in \mathbb{S}_\ell$, one can define the hash function $h_\alpha(m) = h_{id_n}(m_{\alpha(1)} \ldots m_{\alpha(\ell)})$, where h_{id_n} is the above h. Obviously, the concealing property is valid for the whole set $\{h_\alpha | \alpha \in \mathbb{S}_\ell / \mathbb{S}_\omega\}$ of universal hash functions.

(2) Suppose Bob is restricted to perform single-qubit measurements (for a recent version of quantum digital signatures secure against single-qubit measurements, see [46,47]), and the hash function is given as in the previous item 1 of this theorem. Note that upon receiving δ from Alice, Bob knows both π and $\pi' = \tau \circ \pi \circ \tau^{-1}$, but *does not know which of the two was used* by Alice to encrypt the message m into the quantum state given to him.

The bit value $h_i(m) = h_i(\bar{m}_i)$ of the hash $h(m)$ is the parity of the corresponding block $\bar{m}_i$. For each bit of the block $\bar{m}_i$, say the first bit, its value is completely uncorrelated with $h_i(\bar{m}_i)$, unless one knows the values of all other bits: In order for the measurement observable on the first qubit to be a function of the hash value, Bob needs to know the values of all other bits from that block. On the other hand, within the remaining $\omega - 1$ bits, each one is completely uncorrelated with the others, and the hash value $h_i(\bar{m}_i)$ (since the value of the first bit is not known). Therefore, Bob's choice of the single-qubit measurement observable on each of the remaining $\omega - 1$ qubits is independent on the hash value.

Moreover, the fact that the same secret key π is used to encrypt all bits of the message, thus establishing correlations between single-qubit states, cannot help Bob in choosing the optimal cheating observable: Since for each pair (π,π') the states ρ^b_π and $\rho^b_{\pi'}$ (with $b = 0,1$) are not fully distinguishable, and Bob will, unless he chooses to measure the right M_π (which can happen only with probability $1/2$), inevitably obtain wrong results for single bits [48]. Nevertheless, he can know that he obtained wrong bit values only upon performing measurements on *all* qubits of a single block. Therefore, Bob is forced to perform the same measurement on each of the qubits of a single block (with the possible exception of the last one), which is a function of *both* π and π' [and, as shown above, not $h(m)$].

Let that measurement be given by the (orthogonal) observable $M = 0 \cdot P^+ + 1 \cdot P^-$. In order to achieve $\mathrm{Prob}(+; M, \rho^+_{\pi/\pi'}) = 1$ (i.e., recover the bit value 0 encoded in the state $\rho^+_{\pi/\pi'}$), the eigenspace $\mathcal{H}^+$ for the value 0 of the observable M must contain the ranges of $\rho^+_{\pi/\pi'}$, given by the eigenspaces $\mathcal{H}^+_{\pi/\pi'}$ of the observables $M_{\pi/\pi'}$: $\mathcal{H}^+_\pi \cup \mathcal{H}^+_{\pi'} \subseteq \mathcal{H}^+$. As proven in item 2 of previous theorem, $\mathrm{Prob}(\pm; M_{\pi'}, \rho^+_\pi) = \mathrm{Prob}(\pm; M_{\pi'}, \rho^-_\pi) = 1/2$ (and analogously when measuring M_π on $\rho^\pm_{\pi'}$): The eigenspaces of the two observables M_π and $M_{\pi'}$ corresponding to two different outcomes 0 and 1 are not orthogonal to each other, and thus the two projectors P^+ and P^- cannot be orthogonal. As a consequence, for each pair $\pi,\pi' \in \mathbb{K}_n$, the probability q of distinguishing between $\{\rho^+_\pi, \rho^+_{\pi'}\}$ and $\{\rho^-_\pi, \rho^-_{\pi'}\}$ (i.e., inferring the bit value) is strictly *smaller* than 1. Therefore, the probability of recovering the remaining $\omega - 1$ bits, and thus the whole block, is negligible (of the order of $q^{-\omega}$).

The above argument can be extended to few-qubit measurements. Assume Bob can perform two-qubit measurements with four possible outcomes 00,01,10, and 11, used to infer four distinct two-bit messages. Following the above analysis, in order to *perfectly* infer a two-bit message, say 00, the corresponding eigenspace $\mathcal{H}^{00}$ of the Bob's observable must satisfy $\mathcal{H}^{00}_\pi \cup \mathcal{H}^{00}_{\pi'} \subseteq \mathcal{H}^{00}$, where $\mathcal{H}^{00}_\pi = \mathcal{H}^0_\pi \otimes \mathcal{H}^0_\pi$, etc. Again, since the eigenspaces $\mathcal{H}^s_\pi$ and $\mathcal{H}^{s'}_{\pi'}$, corresponding to two different two-bit strings s and s', are not orthogonal to each other, there is a finite optimal (expected) probability q of distinguishing between different two-bit strings which, for large enough blocks, results in negligible cheating probability (Bob can increase the $1/2$ probability of learning the overall message sent by Alice by only a negligible amount).

Note that as the size of a string that can be coherently measured increases (the number L of qubits that can be coherently measured), the eigenspaces $\mathcal{H}^s_\pi$ and $\mathcal{H}^{s'}_{\pi'}$, corresponding to two different L-bit strings s and s', become more and more close to orthogonal. In order to know how long a block should be, such that for a given L Bob's probability to cheat is negligible, i.e., find $\omega = \omega(L)$, one has to solve the equation $[q(L)]^{\omega/L} = \varepsilon(\omega)$, where $q(L)$ is the (expected) probability of distinguishing between two L-bit strings, and $\varepsilon(\omega)$ is a negligible function. A detailed analysis of the relationship between the length ω of the block $\bar{m}$ (and thus the length $\ell = \omega^2$ of the overall message m) and the size L of disposable coherent measurements is a matter of further study.

(3) To finish the argument of the security of Protocol 1 we show that the protocol is unconditionally *oblivious* against a cheating Alice, i.e., there is no strategy for Alice which would enable her to know with certainty if Bob received the message m or not. This is ensured by the last step of the protocol. Notice that since Alice cannot know beforehand which permutation τ will be chosen by Bob, or know which γ Bob chooses in Step 6 of the protocol, she does not know which measurement $M^{\otimes\ell}_\gamma$, with $\gamma \in \mathbb{K}_n$, is performed by Bob in Step 7. Therefore, in order to know if Bob received the message m, she must prepare a state ρ^m that leads to the same answer m regardless of the measurement selected by Bob. Obviously, in order to satisfy the requirement that Bob learns m in 50% of the cases, she sends uniformly at random either ρ^m, in which case she knows with certainty that Bob got the message, or a completely mixed state $(\mathbb{1}/n!)^{\otimes\ell}$, in which case she knows with certainty that Bob does not get the message except with negligible probability. So, if Alice wants to be nonoblivious, she needs to prepare states that give with certainty the same result for every measurement M_π, and are thus invariant, which is the reason for introducing Step 9. If Bob got the correct message, he can recheck that Alice did not use this cheating strategy by performing another measurement of the same kind but with a different π'. Since the state sent by Alice has to be invariant for any measurement, measuring the state with this π' will lead to the same result as the original choice and Bob will thus abort the protocol. ■

IV. CONCLUSIONS

Oblivious transfer is an important primitive for designing cryptographic protocols and secure multiparty computation schemes. In this paper we proposed a polynomial-time quantum protocol for oblivious transfer of a bit string from

Alice to Bob based on the QSCD_{ff} state distinguishability problem. We showed that, assuming QSCD_{ff} to be polynomially hard even for a quantum computer, our protocol is computationally concealing, unconditionally oblivious, and achieves the goal of transferring information with probability close to $1/2$, as long as Bob can perform only few-qubit coherent measurements (while security against most general coherent attacks performed on *all* qubits of the message is left as an open question). The oblivious and the probabilistic transferring properties rely on the laws of quantum mechanics, while the acknowledgment of the message transfer is ensured by the use of hash functions. Note that, in general, one may not need to use hash functions. For example, if the message sent by Alice is of a particular type, say an NP problem (e.g., SAT or some hard optimization problem), then there is no need to encode the message in the hash value: Bob can verify if he received the message by simply checking if it is the solution of the problem. Also if the message sent by Alice is a binary code of some text in a human language, then the verification of the meaningful information serves as acknowledgment.

ACKNOWLEDGMENTS

This work was partially supported, under the CV-Quantum internal project at IT, by FCT Grants No. PEst-OE/EEI/LA0008/2013 and No. UID/EEA/50008/2013 and EU FEDER, namely via the FCT projects ComFormCrypt Grant No. PTDC/EIA-CCO/113033/2009, and CaPri initiative. André Souto also acknowledges FCT postdoctoral Grant No. SFRH/BPD/76231/2011.

[1] C. Crépeau, in *CRYPTO*, Lecture Notes in Computer Science (Springer, Berlin/Heidelberg, 1987), pp. 350–354.
[2] M. O. Rabin, IACR Eprint archive, Harvard University Technical Report 81 talr@watson.ibm.com 12955 received 21 Jun 2005.
[3] S. Even, O. Goldreich, and A. Lempel, Commun. ACM **28**, 637 (1985).
[4] G. Brassard, C. Crepeau, and J. Robert, in *27th Annual Symposium on Foundations of Computer Science* (IEEE, Washington, DC, 1986), pp. 168–173.
[5] C. Crepeau and M. Santha, in *Sequences II*, edited by R. Capocelli, A. Santis, and U. Vaccaro (Springer, New York, 1993), pp. 360–368.
[6] G. Brassard, C. Crepeau, and M. Santha, IEEE Trans. Inf. Theor. **42**, 1769 (1996).
[7] C. Bennett, G. Brassard, C. Crépeau, and M. Skubiszewska, in *Advances in Cryptology*, edited by J. Feigenbaum (Springer, Berlin/Heidelberg, 1992), Vol. 576, pp. 351–366.
[8] P. Shor, SIAM J. Comput. **26**, 1484 (1997).
[9] S. Wiesner, SIGACT News **15**, 78 (1983).
[10] C. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (IEEE Press, New York, 1984), pp. 175–179.
[11] H. K. Lo and H. Chau, Science **283**, 2050 (1999).
[12] P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000).
[13] D. Mayers, J. ACM **48**, 351 (2001).
[14] V. Scarani, H. Bechmann-Pasquinucci, N. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).
[15] N. Paunković, J. Bouda, and P. Mateus, Phys. Rev. A **84**, 062331 (2011).
[16] H. K. Lo and H. F. Chau, Phys. Rev. Lett. **78**, 3410 (1997).
[17] D. Mayers, Phys. Rev. Lett. **78**, 3414 (1997).
[18] A. Kent, Phys. Rev. Lett. **83**, 1447 (1999).
[19] A. Kent, J. Cryptology **18**, 313 (2005).
[20] N. Bouman, S. Fehr, C. Gonzalez-Guillen, and C. Schaffner, in *Theory of Quantum Computation, Communication, and Cryptography*, edited by K. Iwama, Y. Kawano, and M. Murao, Lecture Notes in Computer Science Vol. 7582 (Springer, Berlin/Heidelberg, 2013), pp. 29–44.
[21] S. Wehner, C. Schaffner, and B. M. Terhal, Phys. Rev. Lett. **100**, 220502 (2008).
[22] C. Schaffner, B. M. Terhal, and S. Wehner, Quantum Inf. Comput. **9**, 963 (2011).
[23] R. Koenig, S. Wehner, and J. Wullschleger, IEEE Trans. Inf. Theor. **58**, 1962 (2012).
[24] N. Ng, S. Joshi, C. Ming, C. Kurtsiefer, and S. Wehner, Nat. Commun. **3**, 1326 (2012).
[25] R. Loura, A. J. Almeida, P. S. André, A. N. Pinto, P. Mateus, and N. Paunković, Phys. Rev. A **89**, 052336 (2014).
[26] H. K. Lo, Phys. Rev. A **56**, 1154 (1997).
[27] G. P. He and Z. D. Wang, Phys. Rev. A **73**, 012331 (2006).
[28] G. P. He and Z. D. Wang, Phys. Rev. A **73**, 044304 (2006).
[29] J. Kilian, in *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88 (ACM, New York, 1988), pp. 20–31.
[30] L. Harn and H. Lin, in *Advances in Cryptology*, edited by H. Imai, R. Rivest, and T. Matsumoto, Lecture Notes in Computer Science Vol. 739 (Springer, Berlin/Heidelberg, 1993), pp. 312–320.
[31] D. Mayers, in *CRYPTO*, edited by D. Coppersmith, Lecture Notes in Computer Science Vol. 963 (Springer, Berlin, 1995), pp. 124–135.
[32] R. Cramer, I. Damgård, and U. Maurer, in *Proceedings of the 19th International Conference on Theory and Application of Cryptographic Techniques*, *EUROCRYPT'00* (Springer-Verlag, Berlin/Heidelberg, 2000), pp. 316–334.
[33] Y. Lindell and B. Pinkas, J. Cryptology **25**, 680 (2012).
[34] Y. Lindell and H. Zarosim, in *Theory of Cryptography*, edited by A. Sahai (Springer, Berlin, Heidelberg, 2013), Vol. 7785, pp. 519–538.
[35] M. Ben-Or, M. Horodecki, D. Leung, D. Mayers, and J. Oppenheim, in *Theory of Cryptography*, edited by J. Kilian, Lecture Notes in Computer Science Vol. 3378 (Springer, Berlin/Heidelberg, 2005), pp. 386–406.
[36] R. Renner and R. König, in *Theory of Cryptography*, edited by J. Kilian, Lecture Notes in Computer Science Vol. 3378 (Springer, Berlin/Heidelberg, 2005), pp. 407–425.
[37] A. Kent, Phys. Rev. Lett. **90**, 237901 (2003).
[38] H. Buhrman, M. Christandl, P. Hayden, H. K. Lo, and S. Wehner, Phys. Rev. Lett. **97**, 250501 (2006).
[39] J. Barrett and S. Massar, Phys. Rev. A **70**, 052310 (2004).
[40] A problem is polynomially hard if there is no polynomial time algorithm for solving it.

[41] A. Kawachi, T. Koshiba, H. Nishimura, and T. Yamakami, J. Cryptology **25**, 528 (2012).

[42] Formally, each permutation $\sigma \in \mathbb{S}_n$ is a bijective function over $\{1,\ldots,n\}$.

[43] Note that the problem could be stated in a more compact way by using mixed states $\rho^{\pm} = \sum_{\pi\in\mathbb{K}_n} \rho_{\pi}^{\pm}/|\mathbb{K}_n|$. Then, the Question would go as follows: are ρ^{+} and ρ^{-} computationally indistinguishable? The choice of somewhat redundant presentation, by explicitly noting that permutation π is unknown, is chosen to emphasize the role of π as a secret key in our Oblivious Transfer protocol (see Protocol 1).

[44] J. Carter and M. Wegman, J. Comput. Syst. Sci. **18**, 143 (1979).

[45] In fact, we require $\ell \leqslant c_1 n$ and $n \leqslant c_2 \ell$. To simplify the presentation, we take $n \geqslant \ell$.

[46] V. Dunjko, P. Wallden, and E. Andersson, Phys. Rev. Lett. **112**, 040502 (2014).

[47] R. J. Collins, R. J. Donaldson, V. Dunjko, P. Wallden, P. J. Clarke, E. Andersson, J. Jeffers, and G. S. Buller, Phys. Rev. Lett. **113**, 040502 (2014).

[48] Note that the strategy of randomly choosing the permutation and performing measurements accordingly within a single block, gives Bob a chance that, with probability $1/2$, confirms that the choice was wrong. This way, even if his choice was wrong, Bob will, with probability exponentially close to 1, learn all of m but a few first blocks. Nevertheless, even in the best case scenario (only the first block $\bar{m}_1$ is wrongly decrypted), there are exponentially many (in fact, $2^{\sqrt{\ell}-1}$) possible messages that Alice could have sent.

Noise and measurement errors in a practical two-state quantum bit commitment protocol

Ricardo Loura,[1,2,*] Álvaro J. Almeida,[3,4] Paulo S. André,[3,4] Armando N. Pinto,[4,5] Paulo Mateus,[1,2] and Nikola Paunković[1,2,†]

[1]*SQIG–Instituto de Telecomunicações, IST-TU Lisbon, 1049-001 Lisbon, Portugal*

[2]*Department of Mathematics, IST, Technical University of Lisbon, 1049-001 Lisbon, Portugal*

[3]*Department of Physics, University of Aveiro, 3810-193 Aveiro, Portugal*

[4]*Instituto de Telecomunicações, University of Aveiro, 3810-193 Aveiro, Portugal*

[5]*Department of Electronics, Telecommunications and Informatics, 3810-193 Aveiro, Portugal*

(Received 17 June 2013; revised manuscript received 13 March 2014; published 30 May 2014)

We present a two-state practical quantum bit commitment protocol, the security of which is based on the current technological limitations, namely the nonexistence of either stable long-term quantum memories or nondemolition measurements. For an optical realization of the protocol, we model the errors, which occur due to the noise and equipment (source, fibers, and detectors) imperfections, accumulated during emission, transmission, and measurement of photons. The optical part is modeled as a combination of a depolarizing channel (white noise), unitary evolution (e.g., systematic rotation of the polarization axis of photons), and two other basis-dependent channels, namely the phase- and bit-flip channels. We analyze quantitatively the effects of noise using two common information-theoretic measures of probability distribution distinguishability: the fidelity and the relative entropy. In particular, we discuss the optimal cheating strategy and show that it is always advantageous for a cheating agent to add some amount of white noise—the particular effect not being present in standard quantum security protocols. We also analyze the protocol's security when the use of (im)perfect nondemolition measurements and noisy or bounded quantum memories is allowed. Finally, we discuss errors occurring due to a finite detector efficiency, dark counts, and imperfect single-photon sources, and we show that the effects are the same as those of standard quantum cryptography.

DOI: 10.1103/PhysRevA.89.052336 PACS number(s): 03.67.Dd, 03.67.Hk, 42.50.Ex

I. BIT COMMITMENT

Among security tasks, the bit commitment protocol holds a prominent role as it represents a computational primitive for many important information-processing protocols. It is a two-party protocol that consists of two phases: the *commitment* and the *opening* phases. In the commitment phase, one client (Alice) *commits* to a value of a bit (commits to either 0 or 1) at a certain moment in time t_0. After performing the commitment, Alice finalizes the protocol by revealing (*opening*) her choice to the other client (Bob), at some later moment in time t_1. The commitment to a certain value could be seen, for instance, as a promise to either perform a certain action in a future moment in time $t_2 \geqslant t_1$ (e.g., buy a house from Bob for a given fixed price X)—commitment to 1, or not—commitment to 0. The protocol has to fulfill three security requirements: Alice cannot change her commitment later in time, in particular during the opening phase (the protocol is *binding*); Bob cannot learn Alice's commitment before the opening phase (the protocol is *concealing*); if both clients are honest (if they execute the protocol according to the rules), then Bob will successfully open Alice's commitment (the protocol is *viable*). Commitment schemes are currently an important phase in several cryptographic protocols, in particular in *zero-knowledge proof systems* and *authentication protocols* (for a more detailed description, see Appendix A).

The idea behind the classical solutions to this problem is to *lock* Alice's commitment in a secure "safe" (the commitment phase), such that without the key it is impossible to break into it, and give that "safe" to Bob. During the opening phase, Alice gives Bob the key, and he learns her commitment. One way of doing this is to use ordinary keys and locks. Another way is to perform the locking by encrypting the commitment, using a secret encryption key. In both cases, the solutions have to meet the binding and concealment requirements. Unfortunately, it is not possible to perform a bit commitment protocol that is unconditionally secure. If Alice chooses to protect her commitment by placing it in a real safe, since there are no unbreakable safes, the protocol would not be unconditionally concealing. If she chooses to protect her commitment by encrypting it using a secret key known only by her, she can achieve the concealing requirement, but then the protocol would no longer be binding. In other words, no unbreakable encryption scheme is binding at the same time—whatever commitment value Alice encrypted during the commitment phase, she can always present a suitable key that would decrypt to either of the two commitment values.

Attempts to solve this problem using quantum systems have been made previously [1], but it was shown that no quantum bit commitment protocol can be both unconditionally binding and concealing [2]. Nevertheless, it is possible, using the current technological limitations, to perform a practical quantum bit commitment protocol that will be secure in the foreseeable future [3]. The protocol is based on the famous BB84 quantum cryptographic protocol [4]. The proposed protocol is binding due to the unconditional security of the BB84 protocol [5] and the fact that we do not have long-term stable quantum memories, nor do we have apparatuses able to perform nondemolition measurements of photons. Practical bit commitment protocols whose security is based on a limited amount of quantum memories were studied previously [6], as

*ricardoloura@gmail.com

†npaunkov@math.ist.utl.pt

 ©2014 American Physical Society

well as protocols whose security is based on having imperfect (due to unavoidable noise) quantum memories [7]. One such protocol, using entangled states, was recently reported to be experimentally performed [8]. Finally, we note that the above no-go theorem on unconditional security of (quantum) bit commitment schemes is applicable only to nonrelativistic protocols. Using relativistic effects, it is possible to design an unconditionally secure bit commitment protocol [9], which was recently implemented [10].

In this paper, we present a two-state version of the practical quantum bit commitment protocol based on the B92 cryptographic protocol [11]. Its practical security relies on the fact that long-term stable quantum memories and nondemolition measurements are currently out of reach. Provided those technological limitations, our protocol is more secure than classical counterparts, as its security is based on physical laws rather than on computational assumptions. We also study the effects of noise, source imperfections, and measurement errors on the protocol's security. In particular, we show that adding a certain amount of white noise always increases the chances that a dishonest Alice will cheat Bob and postpone her decision until the opening phase. Finally, we analyze the security of the protocol in the presence of (im)perfect nondemolition measurements and noisy or bounded quantum memories.

II. TWO-STATE QUANTUM PROTOCOL

The protocol is based on quantum complementarity—the impossibility of simultaneously measuring two noncommuting observables. Therefore, one has to decide to measure only one out of two possible observables of a physical system, and obtain information about only one of two features of a system. The choice of measurement can be interpreted as a commitment to a bit value, and the measurement outcome used as a proof of this particular choice (i.e., commitment). This is somewhat the opposite approach to that used in classical solutions: instead of (securely) imprinting the information of a commitment choice into a state of a physical system (writing down an encrypted message on a piece of paper, for example), the choice is done by acquiring information about *only one out of two* possible features of a physical system. This approach has already been used for designing quantum contract signing [12,13] and simultaneous dense coding protocols [14]. In those cases, the security of the protocols is provided by the laws of physics (e.g., quantum mechanics), rather than by the computational complexity of the decryption schemes. Also, a "probabilistic" two-state quantum-bit string commitment protocol, based on the same mechanism, was recently proposed [15], in which Alice commits to a string of n bits, such that Bob can learn not more than $m < n$ bits, up to negligible probability (note that because this protocol is quantum, its unconditional security is not implying the existence of unconditionally secure quantum bit commitment schemes, as would be the case for its classical counterpart [15]). For similar work on coin tossing and bit-string generation, see [16–19].

In addition to the commitment and the opening phases, the quantum protocol begins with the *initialization* phase, during which Bob prepares a number of identical physical two-level systems (qubits) on which Alice is to perform the commitment measurement (the same measurement on each qubit). Bob sends to Alice a number of qubits, each randomly prepared in one of two given quantum states ($|0\rangle$ or $|1\rangle$), without revealing any information about the prepared states. During the commitment phase, Alice chooses only one out of two noncommuting observables, $\hat{C}_0$ or $\hat{C}_1$ (given by the two states used by Bob, see below), measures it on each qubit received from Bob, and keeps the record of measurement outcomes. Finally, during the opening phase, she reveals the results to Bob, which serves as a proof of her commitment.

We require that the qubit states $|0\rangle$ and $|1\rangle$ used in the protocol are not orthogonal,[1] $\langle 0|1\rangle = \cos\theta$, with $\theta \in (0,\pi/2)$. Let us denote the states orthogonal to $|0\rangle$ and $|1\rangle$ as $|0^\perp\rangle$ and $|1^\perp\rangle$, respectively: $\langle 0^\perp|0\rangle = 0$ and $\langle 1^\perp|1\rangle = 0$. This way, we defined two (orthonormal) bases $\mathcal{B}_0 = \{|0\rangle,|0^\perp\rangle\}$ and $\mathcal{B}_1 = \{|1\rangle,|1^\perp\rangle\}$, which in turn define two orthogonal observables $\hat{C}_0$ and $\hat{C}_1$:

$$\begin{aligned} \hat{C}_0 &= 0 \times |0\rangle\langle 0| + 1 \times |0^\perp\rangle\langle 0^\perp|, \\ \hat{C}_1 &= 1 \times |1\rangle\langle 1| + 0 \times |1^\perp\rangle\langle 1^\perp|. \end{aligned} \tag{1}$$

Finally, we list the eigenvectors of $\hat{C}_1$ expressed[2] in the $\mathcal{B}_0$ basis,

$$\begin{aligned} |1\rangle &= \cos\theta|0\rangle + e^{i\phi}\sin\theta|0^\perp\rangle, \\ |1^\perp\rangle &= \sin\theta|0\rangle - e^{i\phi}\cos\theta|0^\perp\rangle, \end{aligned} \tag{2}$$

and vice versa,

$$\begin{aligned} |0\rangle &= \cos\theta|1\rangle + \sin\theta|1^\perp\rangle, \\ |0^\perp\rangle &= e^{-i\phi}(\sin\theta|1\rangle - \cos\theta|1^\perp\rangle) \end{aligned} \tag{3}$$

for some $\phi \in [0,2\pi)$.

We can now give a more detailed description of our two-state practical quantum bit commitment protocol. It consists of three phases, arranged in chronological order ($T_0 < T_1 < T_2$):

The initialization phase. At time T_0, Bob randomly chooses a string of N bits $(b_1,b_2,\ldots,b_N)$, with $b_k \in \{0,1\}$, and sends a string of N qubits to Alice, each prepared in the pure state $|b_k\rangle$, and emitted at time t_k, with $k \in \{1,2,\ldots,N\}$. Bob keeps the information of the states $|b_k\rangle$ of each qubit, as well as the times t_k of the emission of each qubit. We assume that $t_1 < t_2 < \cdots t_N$ and that $t_N - t_1 \ll T_2 - T_1$.

The commitment phase. At time T_1, Alice starts measuring on *all* the qubits received *only one* observable, either $\hat{C}_0$ or $\hat{C}_1$, depending on her commitment choice ($\hat{C}_0$ corresponds to the commitment to value 0, $\hat{C}_1$ to value 1). She announces the

[1]Note that we fix the scalar product to be a real number. As the two states $|0\rangle$ and $|1\rangle$ are used to design Alice's measurement observables $\hat{C}_0$ and $\hat{C}_1$, the relative phase between the two vectors is irrelevant, and for simplicity we take it to be zero. Note also that, as the relevant quantity in the calculations is $\cos\theta$, for reasons of simplicity we do not use a standard Bloch representation of pure qubit states. A simple substitution $\theta \to \theta/2$ in the formulas gives the results presented below in the standard form of a Bloch representation.

[2]Note that, for reasons of simplicity, we take the particular choice of the gauge for state $|1^\perp\rangle$, fixing the relative phases between it and states $|0\rangle$ and $|0^\perp\rangle$ to be 0 and ϕ, respectively.

arrival times of each qubit (which are at the same time the times of measurement of each qubit; see below for the discussion), a string $(\tau_1,\tau_2,\ldots,\tau_n)$, with $\tau_1 = T_1$, and keeps the record of the measurement results to her, a string $(r_1,r_2,\ldots,r_n)$, with[3] $n \leqslant N$.

The opening phase: At time T_2, Alice reveals her commitment $c \in \{0,1\}$ (e.g., the measurement observable $\hat{C}_c$), together with the measurement results r_i, with $i \in \{1,\ldots,n\}$, to Bob.

Note that not all qubits sent by Bob arrive to and are measured by Alice. Thus, $n \leqslant N$, and for each index i labeling Alice's measurement times τ_i and results r_i, there is a corresponding index $k = k(i)$ labeling Bob's qubit emission times $t_{k(i)}$ and corresponding bits $b_{k(i)}$.

First, we discuss in more detail the commitment mechanism. The description of the commitment phase states that measuring $\hat{C}_0$ corresponds to the commitment to value 0, while measuring $\hat{C}_1$ corresponds to the commitment to value 1. From the expression of measuring observables in terms of the states $|0\rangle$ and $|1\rangle$, and the states orthogonal to them, given by Eq. (1), we see that when a bit value b_k, defining the qubit's quantum state $|b_k\rangle$, "coincides" with the measuring observable $\hat{C}_c = \hat{C}_{b_k}$, i.e., $b_k = c$, then the corresponding measurement outcome r_i, for which $k = k(i)$, coincides with the bit value b_k, i.e., $r_i = b_k$. This way, we can interpret the measurement outcome r_i as Alice's inference of Bob's bit value $b_{k(i)}$: if the bit value and the observable "coincide," the inference will be right; otherwise, it will be random.[4] If by $p_c(r|b)$, with $c,r,b \in \{0,1\}$, we denote a conditional probability that a result r is obtained when measuring observable $\hat{C}_c$ on state $|b\rangle$, then the overall conditional probabilities are given by the following expressions:

(i) Alice measures $\hat{C}_0$:

$$\begin{aligned} &p_0(0|0) = 1, \quad p_0(1|0) = 0, \\ &p_0(0|1) = \cos^2\theta, \quad p_0(1|1) = \sin^2\theta. \end{aligned} \tag{4}$$

(ii) Alice measures $\hat{C}_1$:

$$\begin{aligned} &p_1(0|0) = \sin^2\theta, \quad p_1(1|0) = \cos^2\theta, \\ &p_1(0|1) = 0, \quad p_1(1|1) = 1. \end{aligned} \tag{5}$$

Thus, the above conditional probabilities give the signature of the commitment: if Alice measures $\hat{C}_0$, the statistics of her measurement outcomes will be given by (4); otherwise, they will be given by (5). It also serves as the proof of her commitment during the opening phase: only if the statistics of $\{r_i\}$, with respect to $\{b_{k(i)}\}$, are "close enough" to $p_c(r|b)$ did Alice commit to the bit value c. If $n(r|b)$ is the number of measurements performed by Alice on qubits received in the state $|b\rangle$, *with outcome* r, and $n(b)$ is the *total number* of qubits received in the state $|b\rangle$, then define $q(r|b) = \frac{n(r|b)}{n(b)}$. By "close enough" to, say, $p_0(r|b)$ ($c = 0$ commitment), we now mean that the probability $P(q||p_0)$ that the measurement statistics $q(r|b)$ were produced by a random source $p_0(r|b)$ is bigger than a certain threshold value $\alpha > 0$ (Alice did measure $\hat{C}_0$). This represents Bob's criterion to accept Alice's commitment to $c = 0$, and analogously for $c = 1$. Moreover, for the protocol to be viable, we require that the statistics $q(r|b)$ which pass the test $P(q||p_0) > \alpha$ of committing to value $c = 0$ are unlikely to be obtained by measuring $\hat{C}_1$, i.e., we require that

$$P(q||p_0) > \alpha \Rightarrow P(q||p_1) < \beta \tag{6}$$

(and analogously for statistics that pass the test of committing to $c = 1$). The security parameters α and β are to be set by Bob and the protocol designer, respectively, depending on the particular equipment used and the desired level of confidence.

Let us now discuss the protocol's (*practical*) security and show that it is indeed both binding and concealing. The protocol must guarantee that at a certain moment of time T_1, Alice commits to a bit value such that Bob cannot learn her commitment until she reveals it at time T_2, and Alice cannot change her decision after T_1, in particular at T_2.

Alice makes the commitment by measuring one of the two observables, during a period of time $[\tau_1,\tau_n]$. The protocol requires that Alice announces the arrival times of the qubits. To do so, she either performs her measurements straight away (meaning that she commits to a certain value), or she performs a nondemolition measurement, thus detecting the presence of a photon without either destroying it or affecting its polarization state. However, photonic nondemolition measurements are still in their infancy stage (see [20,21]). In [21], for instance, one performs a nondemolition measurement in a controlled environment (photons are in a confined cavity). This type of measurement is thus suitable for computational rather than optical cryptographic purposes. Furthermore, even if Alice had access to a nondemolition measuring apparatus, she would need, in order to perform her measurements later in time, to have access to stable long-term quantum memories. Despite recent efforts, current technology allows only for very short-term noisy memories, usually implemented by fiber optic cable (see Fig. 5 in [8] for an example): adding extra fiber optic cable makes the measurements much more prone to errors, and increases the qubit losses exponentially in the length of the cable. Consequently, Alice performs her measurements as soon as she receives the qubits sent to her by Bob, turning the arrival times of each qubit into the measurement times τ_i as well. Therefore, $\tau_n - \tau_1 \leqslant t_N - t_1 \ll T_2 - T_1$, and for all practical purposes we can say that the commitment in fact occurred at time T_1. Note that in the ideal case when all qubits sent by Bob arrive to Alice and are detected by her, $n = N$, we have $\tau_i = t_{k(i)} + vl$, with v being the speed of the qubits and l the distance between Alice and Bob.

For these same reasons, the protocol is binding: Alice must perform her measurements as the qubits arrive, and thus she must make a commitment at time T_1, and not later. Otherwise,

[3]Note that some (in fact with today's technology, many) qubits sent will be either lost during their transmission or not detected due to imperfect detectors.

[4]We call this *ambiguous inference*: when measuring observable $\hat{C}_0$ and obtaining result 0, we are not sure that the qubit was in the state $|0\rangle$, while obtaining 1 means that the qubit's state was definitely $|1\rangle$ (and analogously for $\hat{C}_1$). In other words, when measuring $\hat{C}_c$, *all* qubits prepared in the state $|c\rangle$ will be among those for which the result was c, with the drawback that some among them will be prepared in the state $|1\rangle$. If we are interested in what we call *unambiguous inference*, i.e., when measuring $\hat{C}_c$ and obtaining c we want to be sure the state was $|c\rangle$, then we simply have to relabel the measuring observables.

she could keep the qubits in a quantum memory and perform her measurements—commit to a value—later in time. On the other hand, as a consequence of the laws of quantum mechanics, there exists no measurement that would provide Alice with the knowledge of the states prepared by Bob for *all* of the received qubits: she cannot know both $p_0(r_i|0)$ and $p_1(r_i|1)$ for every i. Thus, after performing her measurements, Alice cannot pass both the test (4) of committing to the value 0 and the test (5) of committing to the value 1. Note that it is essential that the commitment tests (4) and (5) contain both $p_0(r_i|0)$ and $p_1(r_i|1)$, *as well as* $p_0(r_i|1)$ and $p_1(r_i|0)$, otherwise the protocol would not be binding. Indeed, Alice must be able both to identify states corresponding to her commitment choice *and* to have proper statistics on states not corresponding to her choice. Otherwise, she could trivially pass both tests even without any measurements by simply setting $r_i = 0$ for all i's if she wants to pass the test of committing to the value 0, and $r_i = 1$ otherwise.

Regarding the second security requirement, that of concealment, it is obvious that Alice's measurements reveal no information about her measurement outcomes, and thus of her commitment to a spatially distant Bob. Sending entangled states would obviously not help, due to nonsignaling and causality: Bob cannot infer the choice of Alice's local measurement by measuring his part of the entangled pair, spatially distant from Alice.

III. OPTICAL NOISE

The above discussion of the commitment mechanism, based on quantum complementarity, was done for the ideal case of noiseless channels and perfect sources and measurements.[5] In particular, the expressions (4) and (5) for conditional probabilities are obtained under this assumption. In this and the following section, we will discuss the case of a noisy environment. Regarding a future implementation of the protocol, in which the qubit states are encoded into the polarization of single photons [22–25], we will discuss the case of optical realizations of the protocol. Nevertheless, our theoretical approach could be easily applied, with suitable small modifications, to other types of physical realizations as well. First, we will consider optical noise, while in the next section we will discuss nonoptical effects, such as imperfect single-photon sources, losses during the transmission, and imperfect detectors (detector efficiency and dark counts).

Note that in this protocol, unlike the case of quantum cryptography, we are interested in more than just one quantity. The quantum-bit error rate (QBER), used to study the effects of noise in quantum cryptography [see Eqs. (31)–(33) in [26], p. 166], is the ratio between the wrongly measured versus the total number of qubits received when we measure in the *same basis* in which the qubits were prepared. In our case, though, we are interested in the ratios, i.e., the (conditional) probabilities of both the case of measurement in the same basis and the case of measurement in a basis different from that in which the qubits were prepared. In the ideal case, the conditional probabilities were given by the expressions (4) and (5). In the following, we will present the corresponding conditional probabilities for the cases of depolarizing channel, bit-flip and phase-flip channels, and arbitrary unitary evolution. At the end of this section, we will combine the four contributions into a single one.

[5]The only effects of the environment considered were the fact that not all of the qubits sent will arrive to Alice, and that some of those that arrive will not be detected at all. This was done only qualitatively, so that the protocol could be properly defined (mentioning that $n \leqslant N$).

A. Depolarizing channel

The depolarizing channel is a model of white noise: with probability $(1 - p)$, the state of a system (in our case qubit) stays the same, while with probability p it becomes totally mixed. Note that in this case, the probability to obtain the result corresponding to the initial state is higher than $(1 - p)$: even if, after passing the channel, the state of the system turns out to be totally mixed (which happens with probability p), there is still nonzero probability to obtain the result corresponding to the initial state. In this case, the probability to obtain the "wrong" (e.g., opposite) result, when measuring in the same basis in which the qubits were prepared, is

$$p_0(1|0) = p_1(0|1) = p/2. \tag{7}$$

This is merely the optical part of the QBER, given by Eq. (34) from [26]: $\mathrm{QBER}_{\mathrm{opt}} = (1 - V)/2$. Using this formula, where V is the "visibility" parameter, we get that $V = (1 - p)$, which is precisely the probability that the state will pass the channel intact—hence the term "visibility" (a synonym, in a sense, of "transparency").

The depolarizing channel has no preferred axis of action, in the sense that its action is the same in each basis of the system's Hilbert space. The noise is the same along each axis, and is the most dominant type of noise or error that occurs, as it is a model of white noise. The Kraus decomposition (or the so-called operator-sum representation; see [27], p. 360, Sec. 8.2.3) of the (super)operator representing the depolarizing channel is, for the case of qubit states, given by

$$\mathcal{E}_d(\hat{\rho}) = (1 - p)\hat{\rho} + p\frac{\hat{I}}{2} = \left(1 - \frac{3}{4}\right)\hat{\rho} + \frac{p}{4}\sum_{i=1}^{3}\hat{\sigma}_i\hat{\rho}\hat{\sigma}_i, \tag{8}$$

where $\hat{\rho}$ is a general mixed state representing the initial qubit state, $\hat{I}$ is the identity operator, and $\hat{\sigma}_i$ are the standard Pauli operators. Note that the first equality is the general definition, while the second one is a particular expression for a two-dimensional qubit case.

Using the above definition (8), one obtains the relevant conditional probabilities, analogous to (4) and (5) obtained for the ideal case ($\eta = \mathrm{QBER}$ represents the optical part of the quantum-bit error rate):

(i) Alice measures $\hat{C}_0$:

$$\begin{aligned} &p_0(0|0) = 1 - \frac{p}{2} = 1 - \eta, \quad p_0(1|0) = \frac{p}{2} = \eta, \\ &p_0(0|1) = (1 - p)\cos^2\theta + \frac{p}{2} = (1 - 2\eta)\cos^2\theta + \eta, \\ &p_0(1|1) = (1 - p)\sin^2\theta + \frac{p}{2} = (1 - 2\eta)\sin^2\theta + \eta. \end{aligned} \tag{9}$$

(ii) Alice measures $\hat{C}_1$:

$$p_1(0|0) = (1-p)\sin^2\theta + \frac{p}{2} = (1-2\eta)\sin^2\theta + \eta,$$
$$p_1(1|0) = (1-p)\cos^2\theta + \frac{p}{2} = (1-2\eta)\cos^2\theta + \eta,$$
$$p_1(0|1) = \frac{p}{2} = \eta, \quad p_1(1|1) = 1 - \frac{p}{2} = 1-\eta. \tag{10}$$

B. Bit-flip channel

The other two types of channels, bit- and phase-flip, are basis-dependent. This means that, in the case of a bit-flip in the $\mathcal{B}_0$ basis, it flips (changes) the state $|0\rangle$ into $|0^\perp\rangle$ (and vice versa) with probability p, where, by construction, $\langle 0|0^\perp\rangle = 0$. But if the state is a general superposition $a|0\rangle + b|0^\perp\rangle$, the flipped state, $b|0\rangle + a|0^\perp\rangle$, will not in general be orthogonal to the initial state, $a|0\rangle + b|0^\perp\rangle$, so it will not be a bit-flip in other bases. Therefore, such noise or errors are expected to occur in cases in which we can isolate a preferable axis (and thus a basis), which is the case of a measurement of an observable (or a preparation of a certain state, which is a basis state of a certain observable). The operator-sum representation of the bit-flip channel is

$$\mathcal{E}_{b_0}(\hat{\rho}) = (1-p)\hat{\rho} + p\hat{\sigma}_{x_0}\hat{\rho}\hat{\sigma}_{x_0}. \tag{11}$$

For simplicity, we will not present the relevant conditional probabilities here, since they can be easily derived from the general formula (15) for combined noises, given in Sec. III E.

Note that the above channel flips the basis states $|0\rangle$ into $|0^\perp\rangle$ (and vice versa), such that the matrix representation of the operator $\hat{\sigma}_{x_0}$ in the basis $\mathcal{B}_0 = \{|0\rangle, |0^\perp\rangle\}$ is the Pauli matrix $\sigma_x = [\hat{\sigma}_{x_0}]_{\mathcal{B}_0} = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. We could also consider a bit-flip channel where states $|1\rangle$ and $|1^\perp\rangle$ are flipped, using $\hat{\sigma}_{x_1} = |1\rangle\langle 1^\perp| + |1^\perp\rangle\langle 1|$ instead of $\hat{\sigma}_{x_0}$.

C. Phase-flip channel

The second basis-dependent operation to model the noise is the phase-flip channel. It "flips" the phase of one of the two basis vectors. Below, as in the previous case, we fix the basis $\mathcal{B}_0$ and represent the flip of the phase of $|0^\perp\rangle$ by the operator $\hat{\sigma}_{z_0}$ whose matrix representation is again a Pauli matrix $\sigma_z = [\hat{\sigma}_{z_0}]_{\mathcal{B}_0} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$. The operator-sum representation of the phase-flip channel is

$$\mathcal{E}_{p_0}(\hat{\rho}) = (1-p)\hat{\rho} + p\hat{\sigma}_{z_0}\hat{\rho}\hat{\sigma}_{z_0}. \tag{12}$$

Again, the relevant conditional probabilities are shown in (15). In analogy with the bit-flip channel, here as well we could consider a channel $\mathcal{E}_{p_1}$ given by the operator $\hat{\sigma}_{z_1} = |1\rangle\langle 1| - |1^\perp\rangle\langle 1^\perp|$.

D. Unitary evolution

Finally, the unitary evolution could be used to model the cases for which we have a constant "rotation" of the state of a system. For example, we may send qubits as photons through an optical fiber which, due to bad twisting, rotates the polarization angle by a fixed ratio per unit length [28]. The unitary evolution is given by

$$\mathcal{E}_u(\hat{\rho}) = \hat{U}\hat{\rho}\hat{U}^\dagger, \tag{13}$$

where the arbitrary $U(1)$ unitary operator is, up to an irrelevant global phase, given by its matrix representation (say, in the $\mathcal{B}_0$ basis),

$$[\hat{U}]_{\mathcal{B}_0} = \begin{bmatrix} e^{i\lambda}\cos\alpha & -e^{-i\mu}\sin\alpha \\ e^{i\mu}\sin\alpha & e^{-i\lambda}\cos\alpha \end{bmatrix}, \tag{14}$$

with $\alpha \in [0,\pi/2]$ and $\lambda, \mu \in [0, 2\pi)$ [29]. Note that $\sin^2\alpha = \eta$ is the QBER *in the* $\mathcal{B}_0$ *basis.*

The relevant conditional probabilities are given in (15).

E. Total optical noise accumulated during the emission, transmission, and measurement

The next, and final, step in modeling the optical part of the noise is combining the above four contributions in a single set of formulas for conditional probabilities. Our approach is the following. The whole apparatus consists of three parts: the sender (Bob) performing the state preparation using the source of photons, the transmission environment (transmission through the space, optical fiber, etc.), and the receiver (Alice) performing the measurement using essentially detectors and beam-splitters.

In each of the three parts, we can have some of the above described types of noise. The white noise occurs with particular probabilities p_d^p, p_d^t, and p_d^m (p, t, and m stand for preparation, transmission, and measurement, and d for depolarizing channel), characteristic of the equipment and the environment.

The white noise is a generic type of noise that affects all types of instruments and environments. During the preparation and the measurement, we can also have basis-dependent noises: when preparing the $|0\rangle$ state, we can have a bit-flip and a phase-flip in the basis $\mathcal{B}_0 = \{|0\rangle, |0^\perp\rangle\}$, characteristic for this particular preparation procedure. They occur with the corresponding probabilities p_b and p_p (with b and p standing for the bit and phase, respectively; the upper labels are omitted for simplicity). When preparing the $|1\rangle$ state, bit- and phase-flips occur in the basis $\mathcal{B}_1 = \{|1\rangle, |1^\perp\rangle\}$, with the *same* probabilities p_b and p_p as in the previous case, since the two-state preparation apparatuses are rotated with respect to each other (assuming spatial isotropy). In general, bit- and phase-flips could occur along a general axis during the transmission as well, but this is a highly unlikely scenario as usually the transmission environment (space, optical fiber) has no preferential axes (bases) and thus the noise is not likely to be biased in this manner. Finally, a unitary evolution could occur during the transmission, while it is unlikely to happen in sources and detectors, and is thus ignored in the preparation and measurement phases.

Therefore, the overall state of a qubit after "passing through" the preparation apparatus, transmission environment, and the measurement apparatus, just before the detection, can be obtained by the consecutive application of the following channels:

(i) Depolarizing, bit-flip, and phase-flip channels, each with different probabilities (p_d^p, p_b^p, p_p^p), in the preparation apparatus.

(ii) Depolarizing channel, with probability p_d^t and unitary rotation, during the transmission.

(iii) Depolarizing, bit-flip, and phase-flip channels, each with different probabilities (p_d^m, p_b^m, p_p^m), during the measurements.

In each part of the apparatus (sender, transmission environment, and receiver), different channels model different noises that occur at the same time, which is assured by their commutativity. Indeed, all the commutation relations needed are satisfied: depolarizing, bit-flip, and phase-flip channels commute with each other (a consequence of commutation relations for the Pauli matrices and the particular Kraus representations of the three channels in terms of Pauli matrices; e.g., see [27]). Thus, the order of their application during the preparation and measurement are irrelevant. In particular, we can apply the depolarizing channel occurring during the preparation just before the transmission, and the one occurring during the measurement just after the transmission. The depolarizing channel and unitary evolution commute as well (the white noise is isotropic), so that we can treat the white noise by only one parameter: since $\mathcal{E}_d = \mathcal{E}_d^m \circ \mathcal{E}_d^t \circ \mathcal{E}_d^p$, we have $p_d = p_d^m + p_d^t + p_d^p$. On the other hand, the bit- and phase-flips *do not* commute with the unitary evolution (as no axis-dependent operation commutes with the unitary evolution, in general).

Combining the overall noise in the apparatus, depending on the preparation procedure (preparing either $|0\rangle$ of the $\mathcal{B}_0$ basis or $|1\rangle$ of the $\mathcal{B}_1$ basis), we have four distinct channels:

(i) $\mathcal{E}_{00} = \mathcal{E}_{b_0}^m \circ \mathcal{E}_{p_0}^m \circ \mathcal{E}_d \circ \mathcal{E}_u^t \circ \mathcal{E}_{b_0}^p \circ \mathcal{E}_{p_0}^p$, when measuring $\hat{C}_0$ and preparing $|0\rangle$.

(ii) $\mathcal{E}_{01} = \mathcal{E}_{b_0}^m \circ \mathcal{E}_{p_0}^m \circ \mathcal{E}_d \circ \mathcal{E}_u^t \circ \mathcal{E}_{b_1}^p \circ \mathcal{E}_{p_1}^p$, when measuring $\hat{C}_0$ and preparing $|1\rangle$.

(iii) $\mathcal{E}_{10} = \mathcal{E}_{b_1}^m \circ \mathcal{E}_{p_1}^m \circ \mathcal{E}_d \circ \mathcal{E}_u^t \circ \mathcal{E}_{b_0}^p \circ \mathcal{E}_{p_0}^p$, when measuring $\hat{C}_1$ and preparing $|0\rangle$.

(iv) $\mathcal{E}_{11} = \mathcal{E}_{b_1}^m \circ \mathcal{E}_{p_1}^m \circ \mathcal{E}_d \circ \mathcal{E}_u^t \circ \mathcal{E}_{b_1}^p \circ \mathcal{E}_{p_1}^p$, when measuring $\hat{C}_1$ and preparing $|1\rangle$.

The calculation of the corresponding conditional probabilities is rather lengthy, but quite straightforward. We present the final expressions for the conditional probabilities when Alice measures $\hat{C}_0$:

$$
\begin{aligned}
p_0(0|0) &= \tfrac{1}{2}\big[1 + (1-p_d)\big(1-2p_b^m\big)\big(1-2p_b^p\big)\cos 2\alpha\big],\\
p_0(1|0) &= \tfrac{1}{2}\big[1 - (1-p_d)\big(1-2p_b^m\big)\big(1-2p_b^p\big)\cos 2\alpha\big].\\
p_0(0|1) &= \tfrac{1}{2}\big[1 + (1-p_d)\big(1-2p_b^m\big)\big(1-2p_b^p\big)\cos 2\alpha\cos 2\theta\\
&\quad - (1-p_d)\big(1-2p_b^m\big)\big(1-2p_b^p\big)\\
&\quad \times \sin 2\alpha \sin 2\theta \cos(\phi-\lambda-\mu)\big],\\
p_0(1|1) &= \tfrac{1}{2}\big[1 - (1-p_d)\big(1-2p_b^m\big)\big(1-2p_b^p\big)\cos 2\alpha\cos 2\theta\\
&\quad + (1-p_d)\big(1-2p_b^m\big)\big(1-2p_b^p\big)\\
&\quad \times \sin 2\alpha \sin 2\theta \cos(\phi-\lambda-\mu)\big].
\end{aligned}
\tag{15}
$$

The results for the case when Alice measures $\hat{C}_1$ can be obtained from the above ones by exchanging the labels 0s with 1s in the conditional probabilities, for example $p_0(1|0) = p_1(0|1)$, etc. Note that the depolarizing and bit-flip coefficients occur in the same way in all expressions, as $(1-p_d)(1-2p_b^m)(1-2p_b^p)$. Moreover, the effects of both bit-flips (during the state preparation and during the measurement) have the same form as that of a depolarizing channel. Indeed, by introducing $b = 1-(1-2p_b^m)(1-2p_b^p)$, we obtain the joint depolarizing–bit-flip coefficient in a symmetric form $(1-p_d)(1-b)$. Note that, since $p_b^{p/m} \in [0,1/2]$, we have $b \in [0,1]$; the ranges of the two coefficients coincide.

F. Distinguishability between conditional probabilities corresponding to different commitment choices

Each discrete probability distribution $p(i)$, with $i = 1,\ldots,n$, can be seen as a vector $p = (p_1, p_2, \ldots, p_n)$, whose coordinates p_i are the probabilities, $p_i = p(i)$. Yet there is a more suitable representation as a vector $p = (p_1, p_2, \ldots, p_n)$, where the coordinates p_i are square roots of the probabilities, $p_i = \sqrt{p(i)}$. The motivation for this representation is the following: with the standard scalar product, $pq = (p,q) = \langle p|q\rangle = \sum_i p_i q_i = \sum_i \sqrt{p(i)q(i)}$, all vectors representing probability distributions have unit norm, due to the normalization of probabilities to 1. On the other hand, a way to quantify distinguishability between two probability distributions p and q is given by the *fidelity* $F(p,q) \equiv \sum_i \sqrt{p(i)q(i)}$ (known also as the Bhattacharyya coefficient [30]), which is merely the scalar product we just introduced, $pq = \sum_i \sqrt{p(i)q(i)} = F(p,q)$. The more similar the two probabilities are, the bigger the scalar product is (1 when they are identical); the more different, i.e., distinguishable, they are, the smaller the scalar product is (the most distinguishable being the orthogonal ones).

We can use this measure of the probability distinguishability to study the influence of noise on the protocol's performance. In this and the following sections, we will not consider the effects of a possible unitary rotation during the transmission, as it represents a systematic error that can be compensated. Nevertheless, we hope the above results on the conditional probabilities with the influence of a possible unitary rotation could be useful in detecting and eliminating such a systematic error.

When measuring $\hat{C}_0$, we get two probability distributions, each conditioned by the input state $|0\rangle$ or $|1\rangle$: one is $p_0(*|0) = (\sqrt{p_0(0|0)}, \sqrt{p_0(1|0)})$, the other $p_0(*|1) = (\sqrt{p_0(0|1)}, \sqrt{p_0(1|1)})$. When measuring $\hat{C}_1$, we get other two probability distributions, again each conditioned by the input state $|0\rangle$ or $|1\rangle$: one is $p_1(*|0) = (\sqrt{p_1(0|0)}, \sqrt{p_1(1|0)})$, the other $p_1(*|1) = (\sqrt{p_1(0|1)}, \sqrt{p_1(1|1)})$.

The stronger the noise is, the more the resulting conditional probabilities diverge form the ideal case, given by (4) and (5), approaching a pair of totally balanced conditional probabilities. Thus, we may consider the average fidelity between the corresponding probability distributions for the noiseless case and the case of a noise given by the channel $\mathcal{E}$. If $p_c(r|b)$ and $p_c^{\mathcal{E}}(r|b)$ are the probabilities for the ideal noiseless case and the case with a noise given by the channel $\mathcal{E}$, respectively $(c,r,b \in \{0,1\})$, then the average fidelity between the four probability distributions is

$$
\begin{aligned}
\langle F(\mathcal{E})\rangle = \tfrac{1}{4}[&F(\mathcal{E};\hat{C}_0,|0\rangle) + F(\mathcal{E};\hat{C}_0,|1\rangle)\\
&+ F(\mathcal{E};\hat{C}_1,|0\rangle) + F(\mathcal{E};\hat{C}_1,|1\rangle)],
\end{aligned}
\tag{16}
$$

where the four fidelities between the four pairs of probability distributions, each obtained for the case of Alice measuring $\hat{C}_c$, when the state sent by Bob was $|b\rangle$ are

$$\begin{aligned}
F(\mathcal{E};\hat{C}_0,|0\rangle) &= F\big(p_0(*|0),p_0^{\mathcal{E}}(*|0)\big)\\
&= \big[\sqrt{p_0(0|0)p_0^{\mathcal{E}}(0|0)}+\sqrt{p_0(1|0)p_0^{\mathcal{E}}(1|0)}\big],\\
F(\mathcal{E};\hat{C}_0,|1\rangle) &= F\big(p_0(*|1),p_0^{\mathcal{E}}(*|1)\big)\\
&= \big[\sqrt{p_0(0|1)p_0^{\mathcal{E}}(0|1)}+\sqrt{p_0(1|1)p_0^{\mathcal{E}}(1|1)}\big],\\
F(\mathcal{E};\hat{C}_1,|0\rangle) &= F\big(p_1(*|0),p_1^{\mathcal{E}}(*|0)\big)\\
&= \big[\sqrt{p_1(0|0)p_1^{\mathcal{E}}(0|0)}+\sqrt{p_1(1|0)p_1^{\mathcal{E}}(1|0)}\big],\\
F(\mathcal{E};\hat{C}_1,|1\rangle) &= F\big(p_1(*|1),p_1^{\mathcal{E}}(*|1)\big)\\
&= \big[\sqrt{p_1(0|1)p_1^{\mathcal{E}}(0|1)}+\sqrt{p_1(1|1)p_1^{\mathcal{E}}(1|1)}\big].
\end{aligned} \tag{17}$$

The bigger the above expected fidelity is (the more similar the actual probability distributions are to the ideal noiseless ones), the higher is the protocol's security.

One could also analyze the intrinsic properties of the conditional probabilities $p_c^{\mathcal{E}}(*|b)$, obtained when a noise $\mathcal{E}$ is present (for simplicity, when considering the intrinsic properties of distributions in noisy environments, we drop the superscript $\mathcal{E}$). As mentioned before, when presenting the protocol, each choice of Alice's measurement can serve to infer the qubit state, prepared by Bob. Thus, whatever the observable $\hat{C}_c$ she measures, the corresponding distributions obtained for the case when the prepared state is $|0\rangle$, and when it is $|1\rangle$, should be as distinguishable as possible. The fidelities between these two pairs of probability distributions are

$$\begin{aligned}
F(p_0(*|0),p_0(*|1)) &= \sqrt{p_0(0|0)p_0(0|1)}+\sqrt{p_0(1|0)p_0(1|1)},\\
F(p_1(*|0),p_1(*|1)) &= \sqrt{p_1(0|0)p_0(0|1)}+\sqrt{p_1(1|0)p_0(1|1)}.
\end{aligned} \tag{18}$$

The average fidelity between the probability distributions obtained when sending the state $|0\rangle$ and the state $|1\rangle$ is then

$$\begin{aligned}
\langle F(|0\rangle,|1\rangle)\rangle &= 1/2[F(p_0(*|0),p_0(*|1))\\
&\quad + F(p_1(*|0),p_1(*|1))].
\end{aligned} \tag{19}$$

The smaller this average fidelity is, the more distinguishable the two distributions are, thus the better Alice can infer which state was sent by Bob, and the protocol security is better.

Finally, note that Alice's choice of measurement must produce two rather different conditional probability distributions $p_0(*|b)$ and $p_1(*|b)$. Only then can her commitment be imprinted in the set of her measurement outcomes, so that Bob can learn Alice's commitment during the opening phase, and Alice cannot change her decision (the protocol is binding). The average fidelity between the two sets of probability distributions, obtained when measuring $\hat{C}_0$, and $\hat{C}_1$, respectively, is

$$\begin{aligned}
\langle F(\hat{C}_0,\hat{C}_1)\rangle &= 1/2[F(p_0(*|0),p_1(*|0))\\
&\quad + F(p_0(*|1),p_1(*|1))],
\end{aligned} \tag{20}$$

where

$$\begin{aligned}
F(p_0(*|0),p_1(*|0)) &= \sqrt{p_0(0|0)p_1(0|0)}+\sqrt{p_0(1|0)p_1(1|0)},\\
F(p_0(*|1),p_1(*|1)) &= \sqrt{p_0(0|1)p_1(0|1)}+\sqrt{p_0(1|1)p_1(1|1)}.
\end{aligned} \tag{21}$$

Again, the smaller this expected fidelity is, the more distinguishable the two distributions are, which results in higher security of the protocol: the more distinguishable Alice's actions are, the more secure is her commitment (the choice of her action).

Note that in the noiseless case, we have that $\langle F(|0\rangle,|1\rangle)\rangle = \cos^2\theta$, while $\langle F(\hat{C}_0,\hat{C}_1)\rangle = \sin^2\theta$. Thus, according to the first criterion, the best state distinguishability is, as expected, achieved for $\theta = \pi/2$, while the highest measurement distinguishability is achieved for $\theta = 0$ (again, this is a rather trivial fact when Bob sends qubits in only one state, which corresponds to result 0 when measuring $\hat{C}_0$ and 1 when measuring $\hat{C}_1$—the two observables represent the same physical property, with its outcomes being relabeled). The two opposed security requirements become equal for $\theta = \pi/4$, which matches the optimal value for the angle between the states sent by Bob.

The same happens for noisy channels. In particular, in Fig. 1(a) the graph of $|\langle F(|0\rangle,|1\rangle)\rangle - \langle F(\hat{C}_0,\hat{C}_1)\rangle|$ is plotted as a function of θ and p_d in the case of a depolarizing channel. As shown, the optimal choice of θ is $\theta = \pi/4$, unless $p_d = 1$, in which case the measurement results are completely random, and consequently any θ will yield the same behavior. An analogous phenomenon occurs in the bit-flip and the phase-flip channels (see Fig. 1), in which case complete randomness is achieved by setting $p_b = 1/2$ and $p_p = 1/2$, respectively (note that in this plot we extended the domain of the bit-flip coefficient to [0,1], obtaining the plot symmetric around the value $p_b = 1/2$).

Next we present the effects of noise on the average fidelity $\langle F(\mathcal{E})\rangle$ between noisy and noiseless channels (16) [see Fig. 2(a)] and within the noisy channel itself, $\langle F(|0\rangle,|1\rangle)\rangle$ and $\langle F(\hat{C}_0,\hat{C}_1)\rangle$, given by (19) and (20), respectively, for the optimal choice of $\theta = \pi/4$ (recall that in this case the two are equal) [see Fig. 2(b)]. In the plots, as noted in the previous section, the coefficient b stands for $1-(1-2p_b^p)(1-2p_b^m)$, and it represents the joint effect of the bit-flip channels in the preparation and measurement apparatuses. Both figures show the expected behavior. Adding noise gradually takes the probability distributions from a noiseless case to a completely random case. Also note that the effect of the bit-flip channel is the same as the effect of the depolarizing channel.

The other quantity widely used to measure how different probability distributions are is the *relative entropy* (also known as *Kullback-Leibler divergence* [31]; for a review of the use of relative entropy in the field of quantum information, see [32]). For two probability distributions $\{p_i\}$ and $\{q_i\}$, the relative entropy between the two is given by

$$S(p||q) = \sum_i p_i \ln\frac{p_i}{q_i}. \tag{22}$$

Although not formally a distance—it is not symmetric with respect to its arguments—it can still serve as a measure

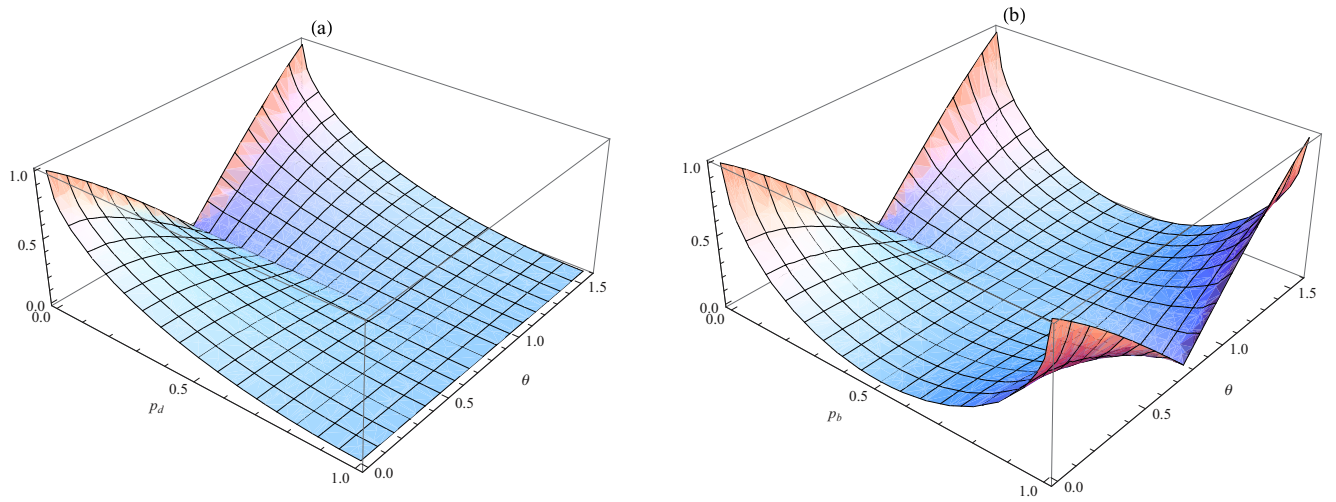


FIG. 1. (Color online) (a) Graph of $|\langle F(|0\rangle,|1\rangle)\rangle - \langle F(\hat{C}_0,\hat{C}_1)\rangle|$ as a function of θ and p_d. (b) Graph of $|\langle F(|0\rangle,|1\rangle)\rangle - \langle F(\hat{C}_0,\hat{C}_1)\rangle|$ as a function of θ and p_b.

of distinguishability.[6] It determines the probability that a random source that emits symbols according to a probability distribution $\{q_i\}$ will produce a sequence of symbols consistent with a source emitting according to $\{p_i\}$ (see *Theorem 4* from [32]).

Here as well we can consider the quantities analogous to those considered in the case of the fidelity, $\langle S(\mathcal{E})\rangle$, $\langle S(|0\rangle||1\rangle)\rangle$, and $\langle S(\hat{C}_0||\hat{C}_1)\rangle$, given by expressions analogous to (16), (19), and (20). Note that, since relative entropy is not symmetric, we can consider six rather than just three quantities. In the case of $\langle S(\mathcal{E})\rangle$, however, only one of the two options is relevant to our study, namely that which quantifies the probability that the noisy environment and imperfect apparatus will reproduce results as in the ideal case given by (4) and (5). We also note that in the noiseless case, the state and measurement distinguishabilities, according to relative entropy, become equal for the optimal value $\theta = \pi/4$. The qualitative results for the relative entropy mimic entirely those for the fidelity, therefore we will not present them here.

[6]Indeed, its infinitesimal form is a true metric, the so-called Fisher information metric (see, for example, [34], p. 87).

IV. OPTIMAL CHEATING STRATEGY FOR ALICE

In this section, we discuss the optimal cheating strategy for Alice if she is allowed to perform only single-qubit measurements. In particular, we analyze the effects of noise on the protocol's security in cases when Alice attempts to cheat. Although the requirement of single-qubit measurements might in general pose a significant constraint, for our practical quantum bit commitment scheme it is rather natural. In other words, not only is it impossible using today's technology to reliably perform large multi-qubit coherent measurements, but in our case Alice would need to have some kind of a stable quantum memory, since Bob sends his qubits sequentially, and at times randomly chosen by him—precisely the equipment

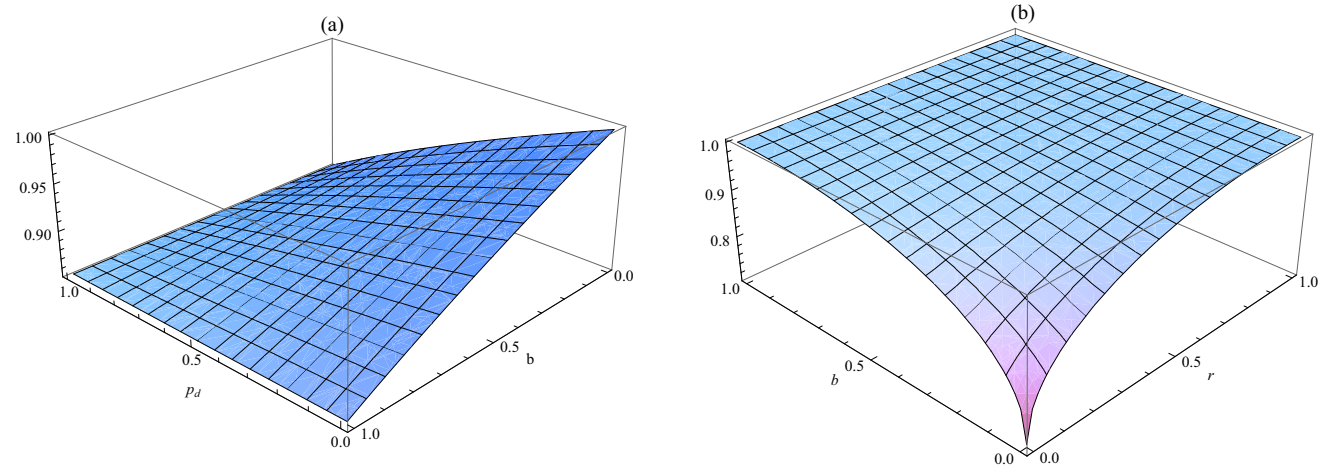


FIG. 2. (Color online) Fidelity $\langle F(\mathcal{E})\rangle$ between noisy and noiseless channels (a), and fidelity $\langle F(|0\rangle,|1\rangle)\rangle = \langle F(\hat{C}_0,\hat{C}_1)\rangle$ within the noisy channel (b), as functions of the depolarizing coefficient p_d and the joint (preparation- and measurement-induced) bit-flip coefficient $b = 1 - (1-2p_b^p)(1-2p_b^m)$.

that is not available today and that makes our (practical) commitment scheme possible.

The goal of a cheating Alice is to break one of the two protocol's security requirements: the binding feature. Alice would like to be able to postpone the moment of her commitment, ideally until the opening phase. To do so, she has to be able to pass both tests of committing to 0 and 1, and since, under the constraints of today's technology, she is forced to perform her measurements immediately upon receiving the qubits from Bob (during the commitment phase), the only option left is to choose a measurement that would provide her with the best possible inference of qubit states sent by Bob. In other words, the optimal measurement has to secure minimal error when discriminating between the two quantum states. This is a well-known problem of ambiguous quantum state discrimination, and the minimal probability of error when discriminating between two general mixed quantum states $\hat{\rho}_0$ and $\hat{\rho}_1$ is given by the famous Helstrom bound [33]:

$$P_e(\hat{\rho}_0,\hat{\rho}_1) = \tfrac{1}{2} + \tfrac{1}{2}\mathrm{Tr}|p_0\hat{\rho}_0 - p_1\hat{\rho}_1|, \tag{23}$$

where p_0 and $p_1 = (1 - p_0)$ are the probabilities of having the state $\hat{\rho}_0$ and $\hat{\rho}_1$, respectively (in our case, $p_0 = p_1 = 1/2$). In the case of pure states and equal *a priori* probabilities, the Helstrom bound is $P_e = (1 - \sin\theta)/2$ and the optimal observable is given by the orthogonal basis vectors $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$, such that (see, for example, [35])

$$|0\rangle = \cos\alpha|\tilde{0}\rangle + \sin\alpha|\tilde{1}\rangle, \quad |1\rangle = \cos\beta|\tilde{0}\rangle + \sin\beta|\tilde{1}\rangle, \tag{24}$$

where $\alpha = \pi/4 - \theta/2$ and $\beta = \pi/4 - \theta/2$. In other words, the basis vectors $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$ are in the plane defined by $|0\rangle$ and $|1\rangle$, and they share the same bisector with them.

For the value of $\theta = \pi/4$, when the protocol's security is maximal, the optimal observable for a cheating Alice is given by the so-called Breidbart basis [36]:

$$\begin{aligned} |\tilde{0}\rangle &= \cos(\pi/8)|0\rangle - \sin(\pi/8)|1\rangle, \\ |\tilde{1}\rangle &= \sin(\pi/8)|0\rangle + \cos(\pi/8)|1\rangle. \end{aligned} \tag{25}$$

To analyze quantitatively the effects of noise on the above cheating strategy, we compare, in analogy with the previous section, how similar various probability distributions are, using the fidelity and relative entropy as distinguishability measures. In particular, we can consider how different the conditional probabilities obtained by a cheating Alice are from those obtained by the honest one. For simplicity, we start by comparing the results obtained by a cheating party, in the presence of noise, with the results of an honest agent, in the ideal noiseless case (4) and (5). We will consider the average fidelity $\langle F(\mathcal{E})\rangle$, given by Eqs. (16) and (17), where in (17) instead of $p_0^{\mathcal{E}}(*|*)$ and $p_1^{\mathcal{E}}(*|*)$, we have the unique cheating probability $p_{\mathrm{ch}}^{\mathcal{E}}(*|*)$. Analogously, we consider the relative entropy $\langle S(\mathcal{E})\rangle$. The results for the fidelity and relative entropy are given in Fig. 3 (note that in the rest of this section, we consider the optimal choice of $\theta = \pi/4$). We observe the qualitative difference between the behavior of the fidelity and the relative entropy: indeed, one can easily see that the entropy decreases slightly with the introduction of small noises (in fact, the optimal value of added noise is rather significant, being slightly over 0.29). This behavior can be taken advantage of in the more realistic situation of both parties being subjected to noise. In fact, as Fig. 4 shows, independently of the channel's noise factor p_d, it is always advantageous for a dishonest party to introduce a small extra noise factor Δp_d, as it decreases the relative entropy between the underlying probability distributions (note that $\Delta p_d \leqslant 1 - p_d$). One can easily prove that the optimal amount of noise $\Delta\tilde{p}_d$ a cheating party should introduce is

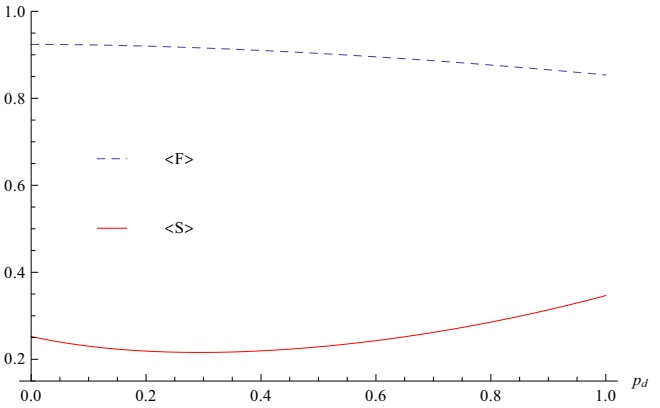


FIG. 3. (Color online) Fidelity $\langle F(\mathcal{E})\rangle$ [dashed (blue) line] and relative entropy $\langle S(\mathcal{E})\rangle$ [full (red) line] between an honest strategy without the presence of noise, and the optimal cheating strategy in the presence of white noise, as a function of the noise parameter p_d.

$$\Delta\tilde{p}_d = \left(1 - \frac{1}{\sqrt{2}}\right)(1 - p_d). \tag{26}$$

The above comparative study of the two distinguishability measures shows two rather conflicting results: while according to the fidelity, noise degrades the chances of a cheating party, according to the relative entropy, it is always advantageous to add a little noise in order to increase the chances that a cheating party will go on unnoticed. Unlike the fidelity, the relative entropy gives the probability that a cheating strategy will produce the distribution of measurement results consistent with that produced by an honest party.

It is interesting to analyze the reasons for such behavior of the relative entropy, and why it is not present in the case of the fidelity. For simplicity, we will analyze only the case

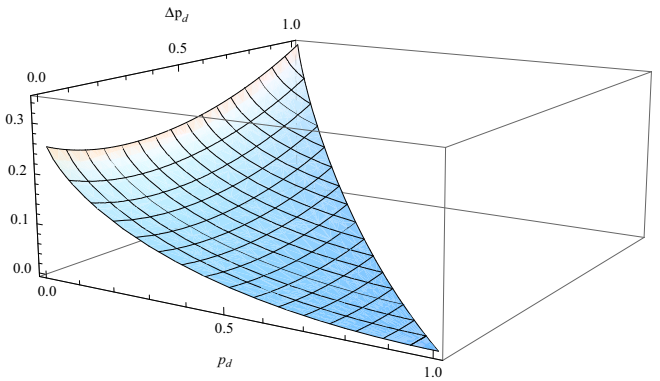


FIG. 4. (Color online) Introducing a small noise Δp_d helps a dishonest party.

when an honest strategy is executed in noiseless circumstances; the general case when both honest and cheating agents are subjected to noise is straightforward. First, we note that unlike the standard quantum cryptographic protocols, such as BB84 [4] and B92 [11], where the only relevant results are those obtained for the cases when the basis of the states sent by Bob and the basis of the observable measured by Alice coincide, here we are interested in the results when the two are not the same. In other words, we are interested not only in how similar the pairs $(p_0(*|0), p^{\mathcal{E}}_{\rm ch}(*|0))$ and $(p_1(*|1), p^{\mathcal{E}}_{\rm ch}(*|1))$ are, but also in the distinguishability of the cross terms given by the pairs $(p_0(*|1), p^{\mathcal{E}}_{\rm ch}(*|1))$ and $(p_1(*|0), p^{\mathcal{E}}_{\rm ch}(*|1))$ of the probability distributions of the measurement results. And it is precisely the cross terms that make a difference: the honest party probability distributions $p_0(*|1)$ and $p_1(*|0)$ are equal and are actually a uniformly random distribution, while the cheating probability distributions $p^{\mathcal{E}}_{\rm ch}(*|0)$ and $p^{\mathcal{E}}_{\rm ch}(*|1)$ are biased, and approach a uniformly random distribution when the noise increases. This means that both the increase in fidelity and the decrease in relative entropy will occur between the pairs of the cross terms for certain range of the noise parameter p_d. Nevertheless, these contributions will affect differently the overall average fidelity $\langle F(\mathcal{E})\rangle$ and the average relative entropy $\langle S(\mathcal{E})\rangle$: the former will always decrease with p_d, while the latter will experience a decrease for a rather broad range of values of p_d. Mathematically, this is just an effect of scaling: fidelity uses a linear scale, and the cross terms are not powerful enough to overcome the behavior of noncross terms, whereas the relative entropy uses a logarithmic scale and allows the cross terms to express themselves better.

V. THE SECURITY OF THE PROTOCOL—A MORE DETAILED ANALYSIS

In Sec. II, we presented a general description of Bob's decision process: in order to accept Alice's commitment to, say, value 0, the probability that the statistics q, formed by the data communicated by Alice, were obtained by measuring $\hat{C}_0$ must be bigger than a certain threshold value α. In addition to that, the probability that the statistics q, which passed Bob's test of committing to 0, were obtained by measuring $\hat{C}_1$ should be smaller than β (the protocol is viable). After analyzing the effects of noise and imperfect photon sources and measurement apparatuses, as well as Alice's optimal cheating strategy, we can study in more detail Bob's criteria for deciding if the results obtained from Alice confirm that she committed to 0 or to 1, or that the results show that she tried to cheat. We can also check the minimum number of photons that need to be measured by Alice for Bob's decisions not to be compromised by Alice's eventual attempt to cheat.

For simplicity, we will analyze the criterion for deciding if given measurement results confirm that Alice committed to 0. The total number of measurement outcomes is $n = n(0) + n(1)$, where $n(0)$ is the number of measurements on photons sent in the state $|0\rangle$, and analogously for $n(1)$. Furthermore, $n(0|0)$ is the number of outcomes 0, when the state $|0\rangle$ is sent, while $n(1|0)$ is the number of outcomes 1 [and analogously for $n(*|1)$]. This way, we have two (conditional) probability distributions:

$$\begin{aligned} q(*|0) &= \left\{ q(0|0) = \frac{n(0|0)}{n(0)}, \quad q(1|0) = \frac{n(1|0)}{n(0)} \right\}, \\ q(*|1) &= \left\{ q(0|1) = \frac{n(0|1)}{n(1)}, \quad q(1|1) = \frac{n(1|1)}{n(1)} \right\}. \end{aligned} \tag{27}$$

To analyze the protocol's security against Alice's attempt to cheat, we introduce a criterion similar to that given by (6):

$$P(q||p_0) > \alpha \Rightarrow P(q||p_{\rm ch}) < \beta. \tag{28}$$

Note that it is more likely to produce statistics that look as if they were obtained by committing to 0 by measuring $\hat{C}_{\rm ch}$ than by measuring $\hat{C}_1$. Therefore, we will only consider the criterion that the statistics $q = \{q(*|0), q(*|1)\}$ were obtained by measuring $\hat{C}_0$, and not by measuring $\hat{C}_{\rm ch}$. Thus, if statistics q pass Bob's test of committing to 0,

$$P(q(*|0)||p_0(*|0)) > \alpha \quad \text{and} \quad P(q(*|1)||p_0(*|1)) > \alpha', \tag{29}$$

then for the protocol to be secure (and consequently viable), we require that

$$P(q(*|0)||p_{\rm ch}(*|0)) < \beta \quad \text{and} \quad P(q(*|1)||p_{\rm ch}(*|1)) < \beta'. \tag{30}$$

(Note that, in general, the thresholds α and α', as well as β and β', need not be equal.)

The above probabilities are given by a simple binomial distribution[7] $\mathcal{B}(n_0; n, p_0)$ [for simplicity, here we define $n = n(0)$, $n_0 = n(0|1)$, and $p_0 = p_0(0|1)$]:

$$P(q(*|1)||p_0(*|1)) = \binom{n}{n_0} p_0^{n_0} (1 - p_0)^{(n-n_0)}, \tag{31}$$

and analogously for other cases. For all practical purposes, the above binomial distribution will behave as a normal distribution $\mathcal{N}(x; \mu_0, \sigma_0)$, with its mean μ_0 and standard deviation σ_0. Thus, we can set parameter α to be $\alpha_{2\sigma} = \Phi(\mu_0 + 2\sigma_0; \mu_0, \sigma_0)$, where $\Phi(x; \mu_0, \sigma_0)$ is the cumulative distribution function of the normal distribution $\mathcal{N}(x; \mu_0, \sigma_0)$. For such α, the statistics obtained by measuring $\hat{C}_0$ will pass the test of committing to 0 in about 97.7% of the cases. Analogously, the distribution defining the second probability in (30) is given by $\mu_{\rm ch}$ and $\sigma_{\rm ch}$, and we can define β to be $\beta_{2\sigma} = \Phi(\mu_{\rm ch} - 2\sigma_{\rm ch}; \mu_{\rm ch}, \sigma_{\rm ch})$, for which the statistics obtained by measuring $\hat{C}_{\rm ch}$ will lead to a successful cheat in only 2.3% of the cases.

It turns out that the second conditions in both (29) and (30) are, albeit qualitatively the same, quantitatively stronger than the first pair of conditions. In Fig. 5, we plot $P(q(*|1)||p_0(*|1))$ (left bump) and $P(q(*|1)||p_{\rm ch}(*|1))$ (right bump) as functions of a depolarizing coefficient p_d and $q_0 = \frac{n(0|1)}{n(0)}$ for $n(0) = 50$

[7] The probability that the distribution $\{p_0, 1 - p_0\}$ yields statistics q [that is, the probability that $\{p_0, 1 - p_0\}$ yields n_0 0s and $(n - n_0)$ 1s] is given by $p_0^{n_0}(1 - p_0)^{(n-n_0)}$ times the number of possible sequences of n_0 0s and $(n - n_0)$ 1s, i.e., $\binom{n}{n_0}$.

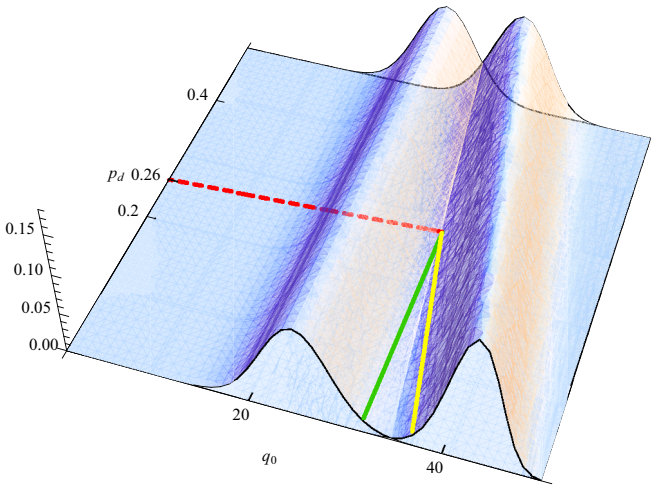


FIG. 5. (Color online) $P(q(*|1)||p_0(*|1))$ (left bump) and $P(q(*|1)||p_{\rm ch}(*|1))$ (right bump), as functions of a depolarizing coefficient p_d and $q_0 = \frac{n(0|1)}{n(0)}$ for $n(0) = 50$. The green (left) full line represents the intersection between the left bump and the $q_0 = \mu_0 + 2\sigma_0$ plane, while the yellow (right) one represents the intersection between the right bump and the $q_0 = \mu_{\rm ch} - 2\sigma_{\rm ch}$ plane, both as functions of p_d.

(i.e., the total number of measurements is about $n = 100$). The second condition in (30) now translates to

$$\mu_0 + 2\sigma_0 \leqslant \mu_{\rm ch} - 2\sigma_{\rm ch}. \tag{32}$$

The latter is satisfied as long as the depolarizing coefficient p_d is smaller than about 26%, which is represented by the intersection of the two full lines in Fig. 5.

We see that already for $n = 100$ we obtain statistics such that the probability to cheat becomes negligible for realistic amounts of noise. Note that this number is smaller than the number of photons needed to be measured in standard quantum key distributions, where a number of results are used not for establishing a secret key, but "wasted" for checking if the communication between Alice and Bob was eavesdropped.

The conditions determining the probabilities α and β in (30), given by (32), can be either strengthened or loosened according to one's needs. One can, for instance, increase the viability of the protocol by setting α to be $\alpha_{3\sigma} = \Phi(\mu_0 + 3\sigma_0; \mu_0, \sigma_0)$, in which case an honest party would successfully pass Bob's test in 99.86% of the cases. On the other hand, one could decrease the security parameter β to $\beta_\sigma = \Phi(\mu_{\rm ch} - \sigma_{\rm ch}; \mu_{\rm ch}, \sigma_{\rm ch})$, which would still allow Bob to spot cheating in about 84.2% of the cases. In Table I, we present various possibilities for p_d^*, the maximum value of p_d for which conditions (29) and (30) are satisfied, as functions of the security parameters α and β [i.e., the conditions analogous to (32)].

TABLE I. Maximum value of p_d as a function of the security parameters α and β.

α	β	p_d^*
$\alpha_{2\sigma} = 97.7\%$	$\beta_{2\sigma} = 2.3\%$	0.26
$\alpha_{3\sigma} = 99.86\%$	$\beta_\sigma = 15.8\%$	0.23
$\alpha_{3\sigma} = 99.86\%$	$\beta_{2\sigma} = 2.3\%$	0.09
$\alpha_{2\sigma} = 97.7\%$	$\beta_\sigma = 15.8\%$	0.42

VI. (IM)PERFECT NONDEMOLITION MEASUREMENTS AND NOISY OR BOUNDED MEMORY

In this section, we analyze the protocol's security under the more realistic assumptions of using finite efficiency nondemolition measurements and a noisy or bounded memory. First, we discuss the use of nondemolition measurements. The ideal nondemolition photon measurement would allow Alice to obtain the photon arrival times without actually destroying them, thus permitting her to keep the qubits in a (noisy) memory and postpone her commitment. Typically, such nondemolition measurements would alter the photon's state of polarization, a contribution that is equivalent to one coming from an imperfect memory and with which we will deal later on. In the case of a finite efficiency, say $p_{nd} < 1$, a certain fraction of arrived photons will not be possible to store. If Alice's equipment simply absorbs such photons, and does not allow measuring the photon polarizations, Bob can detect such cheating attempt by comparing the expected and presented number of results (Bob knows the specifications of the setup, in particular the rate of photon emission, the absorption coefficient of the environment, and the detectors' efficiencies, and can therefore estimate the expected number n of results provided by Alice). If she has an apparatus that can measure the polarization of the absorbed photons, Alice's best strategy is to perform the polarization measurement of the cheating observable $\hat{C}_{\rm ch}$ on the said fraction of $(1 - p_{nd})n$ photons. But this is equivalent to the case of performing an ideal nondemolition measurement, and having a bounded noisy memory, such that only the fraction of $\nu = p_{nd}$ of results are obtained by the stored photons, while the rest are obtained by measuring $\hat{C}_{\rm ch}$. Therefore, in the rest of the section we assume the ideal case of $p_{nd} = 1$.

Regarding quantum memories, we will first again discuss the best-case scenario for a cheating Alice: unrestricted amount of noisy quantum storage. Storing qubits in a memory for a certain "delay time" Δt allows Alice to postpone her commitment and thus break the binding security condition. Since the memory is not ideal, during that time the photon's polarization states will further decohere with the environment, hence decreasing Alice's probability to pass Bob's test (29). We model the noise by a depolarizing quantum channel given by $p_d(\Delta t)$. This way, when measuring, say $\hat{C}_0$, the cheating *a priori* conditional probability distributions $p_0^{\Delta t}(*|*)$, given by $p_d + p_d(\Delta t)$, will differ from those expected by Bob, given by only p_d. As in the preceding section, the threshold value for $p_d(\Delta t)$, for which cheating is not possible, is given by (Bob's decision criterion)

$$P(q(*|0)||p_0(*|0)) > \alpha \quad \text{and} \quad P(q(*|1)||p_0(*|1)) > \alpha, \tag{33}$$

and (security criterion)

$$\begin{aligned} &P(q(*|0)||p_0^{\Delta t}(*|0)) < \beta^{\Delta t} \quad \text{and} \\ &P(q(*|1)||p_0^{\Delta t}(*|1)) < \beta^{\Delta t}. \end{aligned} \tag{34}$$

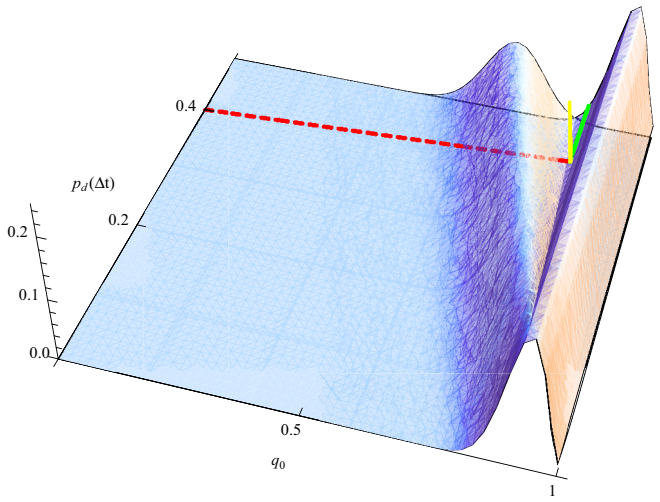


FIG. 6. (Color online) $P(q(*|0)||p_0(*|0))$ (right bump) and $P(q(*|0)||p_0^{\Delta t}(*|0))$ (left bump) as a function of $p_d(\Delta t)$ and $q_0 = \frac{n(0|1)}{n(0)}$ for a fixed value of $p_d = 0.15$ and $n(0) = 50$. The green (right) full line represents the intersection between the right bump and the $q_0 = \mu_0 - 2\sigma_0$ plane, while the yellow (left) one represents the intersection between the left bump and the $q_0 = \mu_0^{\Delta t} + 2\sigma_0^{\Delta t}$ plane, both as functions of $p_d(\Delta t)$.

Note that α is given in terms of μ and σ of the "honest" conditional probability distributions $p_0(*|*)$, while $\beta^{\Delta t}$ is obtained for $\mu^{\Delta t}$ and $\sigma^{\Delta t}$ of the "cheating" conditional probability distributions $p_0^{\Delta t}(*|*)$.

Again, we analyze the case of committing to 0, and for simplicity we choose only the $q(*|0)$ results. Further, in analogy with the previous section, we (re)define $n = n(0)$, $n_0 = n(0|0)$, $p_0 = p_0(0|0)$, and $q(*|0) = \{q = n_0/n, 1 - q = (1 - n_0)/n\}$. This change of notation is due to the fact that $p_0^{\Delta t}(*|1) = p_0(*|1)$ is uniformly random, and independent of Δt. In Fig. 6, we plot $P(q(*|0)||p_0(*|0))$ (right bump) and $P(q(*|0)||p_0^{\Delta t}(*|0))$ (left bump) as functions of $p_d(\Delta t)$, for a fixed value of $p_d = 0.15$. For higher (lower) values of p_d, the corresponding plot is a simple translation to the left (right), along with a rescaling.

Finding the threshold value $p_d^*(\Delta t)$ for $p_d(\Delta t)$ (i.e., for how long can Alice postpone her commitment), as a function of p_d, is now equivalent to solving

$$\mu_0^{\Delta t} + 2\sigma_0^{\Delta t} \leqslant \mu_0 - 2\sigma_0, \tag{35}$$

just as in the previous section [see Eq. (32)]. Again, we can vary the security parameters $\alpha^{\Delta t}$ and $\beta^{\Delta t}$. Various possibilities are presented in Table II.

TABLE II. Maximum value of $p_d(\Delta t)$ as a function of the security parameters α and β.

$\alpha^{\Delta t}$	$\beta^{\Delta t}$	$p_d^*(\Delta t)$
$\alpha_{2\sigma} = 97.7\%$	$\beta_{2\sigma} = 2.3\%$	0.4
$\alpha_{3\sigma} = 99.86\%$	$\beta_{\sigma} = 15.8\%$	0.35
$\alpha_{3\sigma} = 99.86\%$	$\beta_{2\sigma} = 2.3\%$	0.49
$\alpha_{2\sigma} = 97.7\%$	$\beta_{\sigma} = 15.8\%$	0.26

Finally, we briefly discuss the general case of bounded and noisy quantum memories. Let ν be the fraction of the total results n which were obtained by measuring the photons stored in a noisy quantum memory. The optimal strategy for a cheating Alice would then be to measure $\hat{C}_0$ on νn photons, and the cheating observable $\hat{C}_{\text{ch}}$ on the rest of $(1 - \nu)n$ photons that she cannot store. Due to the law of large numbers, the a priori probability of such cheating strategy would be

$$\bar{p}_0(*|0) = \nu p_0(*|0) + (1 - \nu)p_{\text{ch}}(*|0), \tag{36}$$

and analogously for $\bar{p}_0(*|0)$. The security criterion would then be

$$P(q(*|0)||\bar{p}_0(*|0)) < \bar{\beta} \quad \text{and} \quad P(q(*|1)||\bar{p}_0(*|1)) < \bar{\beta}', \tag{37}$$

which can be easily analyzed using the previous results regarding the cheating observable and the noisy memory.

VII. NONOPTICAL DETECTOR ERRORS

In this section, we briefly discuss the effects of nonoptical errors, caused by the imperfect single-photon sources, transmission losses, and imperfect detectors (finite efficiency and dark counts). As the causes of these errors are basis-independent, they will manifest equally as in the case of standard quantum cryptography, when the state $|b\rangle$ sent by Bob and the observable $\hat{C}_c$ measured by Alice coincide, $c = b$. In Appendix B, following [26], p. 166, we present the explicit expressions for nonoptical quantum-bit corrections $\delta p_c(*|b)$, for $c = b$, to the total probabilities $\tilde{p}_c(*|b) = p_c(*|b) + \delta p_c(*|b)$, where $p_c(*|b)$ represents the optical contribution discussed in Sec. III.

As in the case of calculating the nonoptical part of QBER, in cases of measuring a "wrong" observable, when $c \neq b$, we neglect all less probable cases and calculate the error due to dark counts only when no photons arrived at the measuring apparatus (that consists, among other things, of two detectors D_0 and D_1, corresponding to two possible outcomes 0 and 1, respectively). Therefore, we can safely estimate that the number of dark counts in both detectors is equal. In the case of $|\langle 0|1\rangle| = \cos \pi/4 = 1/\sqrt{2}$, when the protocol's security is optimal, the error due to dark counts is thus zero.

VIII. CONCLUSIONS

We presented a two-state practical quantum bit commitment protocol. We discussed the effects of both optical and nonoptical noise. In the latter case, we showed that finite detector efficiency, dark counts, imperfect single-photon sources, and transmission losses have essentially the same effects as in the case of standard quantum cryptography. To quantitatively analyze the effects of the optical part of the noise, we used the fidelity and the relative entropy, two information-theoretic measures of probability distribution distinguishability. As a corollary to our study, using the two distinguishability measures only, we obtained the well-known result that the optimal value of the angle θ between the two quantum states used in the protocol is $\theta = \pi/4$. We also showed a somewhat counterintuitive result that adding a certain amount of white noise can always help a cheating

Alice to postpone her commitment until the opening phase. This effect is a result of the comparison of the results of measurements in cases when the measurement basis does not coincide with the basis from which the state is sent. Although it can be seen by looking at the behavior of both the fidelity and the relative entropy, when averaging over all the possible cases, only the expected relative entropy retains the signature of this effect. Finally, we analyzed the protocol's security when Alice has access to (im)perfect nondemolition measurements and noisy or bounded memories, and we showed that the protocol is robust against such possible attacks.

ACKNOWLEDGMENTS

R.L. thanks Fundação para a Ciência e a Tecnologia (FCT) for support through Ph.D. Grant No. SFRH/BD/79571/2011. A.A. thanks the FCT for support through Ph.D. Grant No. SFRH/BD/79482/2011. A.A., P.A., and A.P. gratefully acknowledge the project PTDC/EEA-TEL/103402/2008 (QuantPrivTel); the FCT and the Instituto de Telecomunicações (IT) under the PEst-OE/EEI/LA0008/2013 program, project "P-Quantum," and the Conselho de Reitores das Universidades Portuguesas (CRUP) project "Ação Integrada E 91/12." P.M. and N.P. gratefully acknowledge the project of SQIG at IT, funded by FCT and EU FEDER projects PTDC/EIA/67661/2006 (QSec), PTDC/EEATEL/103402/2008 (QuantPrivTel), FCT PEst-OE/EEI/LA0008/2013, and IT Projects QuantTel and "P-Quantum," as well as Network of Excellence, Euro-NF. P.M. also gratefully acknowledges ComFormCrypt PTDC/EIACCO/113033/2009.

APPENDIX A

In this appendix, we present a simple example of the application of bit commitment to authentication protocols based on zero-knowledge proof systems. Suppose Alice wants to authenticate herself to Bob by proving to Bob that she knows a solution to a difficult mathematical problem, without actually revealing the solution (thus the name zero-knowledge proof). This requirement is crucial: the knowledge Alice has is unique to her (the problem is difficult, so others cannot solve it *in real time*), and is used as a means of identification. If she discloses it to Bob, he can in the future falsely present himself as Alice, which she would like to prevent.

Consider the following mathematical problem (the so-called *coloring problem*): given a graph $G = (V, E)$, where V is the set of vertices and $E \subseteq V \times V$ the set of edges, and three colors $\{R, Y, B\}$ (red, yellow, and blue), find a coloring $C : V \to \{R, Y, B\}$ such that no two adjacent vertices have the same color, $(u, v) \in E \Rightarrow C(u) \neq C(v)$. This is known to be a hard problem, in fact it is an NP-complete problem (see, for example, [37], p. 1019): if only a graph is given, finding a proper coloring using today's best algorithms requires exponential time with respect to the graphs' complexity. Therefore, for all practical purposes it is safe to assume that Alice is the only person who knows a coloring, and therefore she can use it as her personal identifier.

The way to prove to Bob that she indeed knows the coloring C, without actually revealing this information, is the following. The protocol is probabilistic, consisting of n steps, such that the probability that Alice cheats (convinces Bob she knows the coloring, without actually knowing C) approaches zero exponentially fast with respect to the number of steps n. Each step consists of three consecutive parts:

(i) Alice randomly chooses a permutation $\pi : \{R, Y, B\} \to \{R, Y, B\}$, sets a new coloring $C' = \pi \circ C$, and *commits* to it: she writes down on a piece of paper the colors, according to new coloring C', of all the vertices of a (publicly known) graph G, *locks* it in a secure "safe," keeps the key with her, and gives the "safe" to Bob. She can do so by committing to a string of $2N$ bits, where N is the number of vertices of G: assuming Alice and Bob agreed prior to the protocol on a particular enumeration of the graph's vertices, each ith pair of bits, with $i = 1, \ldots N$, defines the color of the ith vertex (obviously, this is not an optimal encryption); this way, each bit is locked in a different "safe," for which a different key is produced.

(ii) Bob chooses an edge $(u.v)$ and challenges Alice to show him their respective colors.

(iii) Alice *opens* the values of the bit pairs corresponding to the vertex u and the vertex v (gives the keys for the corresponding bits), thus disclosing to Bob the colors $C'(u)$ and $C'(v)$. If they are the same, Alice failed to pass the test and Bob terminates the procedure. Otherwise, they repeat the procedure until Bob is satisfied.

Obviously, Alice can pass the above test only if she indeed knows the coloring C of the graph G. Otherwise, she can only try to partially color the graph (properly, so that the adjacent vertices have different colors), hoping that Bob will not choose vertices that she colored with the same color. If the probability to pass the test in a single step of the protocol is $p < 1$, then the probability to pass the test goes to zero exponentially fast with the number of steps n of the protocol, as $1 - p^n$. Note that although in each step of the protocol Bob learns a coloring of one pair of vertices, after n steps he still did not learn the coloring of n vertices, as in each step Alice chooses different coloring $C' = \pi \circ C$, given by a permutation π, unknown to Bob.

Note the essential importance that Alice's commitment is binding—otherwise, she could, upon learning Bob's choice of vertices u and v, change her commitment and choose the two colors to be different. Also, it is important that the protocol is concealing—otherwise, Bob would be able to learn a coloring of graph G, and thus, in the future, impersonate Alice.

This concept of a cryptographic commitment can be traced back to the early 1980s, with the works of Shamir, Rivest, and Adleman [38], along with those of Blum [39] and finally of Even [40], where the concept was first named. Nowadays, it has found its way into several protocols of many diverse natures, such as e-voting protocols [41], the TESLA authentication protocol [42], and the Schnorr protocol [43], upon which part of Microsoft's U-prove system is based [44].

APPENDIX B

In this appendix, we evaluate the nonoptical contributions $\delta p_c(*|b)$ for $c = b$. As noted when discussing the depolarizing channel, the two out of four probabilities are merely the quantum-bit error rate, $\mathrm{QBER} = \delta p_0(1|0) = \delta p_1(0|1)$, while the other two are then straightforward to obtain,

$\delta p_0(0|0) = -\delta p_0(1|0)$ and $\delta p_1(1|1) = -\delta p_1(0|1)$ [note that, by definition, $\tilde{p}_0(0|0) + \tilde{p}_0(1|0) = 1$ and $p_0(0|0) + p_0(1|0) = 1$]. For example, let Alice measure $\hat{C}_0$. Then, we are interested in cases when result 1 is obtained for qubits in state $|0\rangle$. Let N_{tot} be the total number of qubits received in the state $|0\rangle$, and let N_{wrong} be the number of qubits received in the state $|0\rangle$ for which the wrong result 1 is obtained, during the time interval T. Then, the QBER is given by

$$\text{QBER} = \frac{N_{\text{wrong}}}{N_{\text{tot}}} = \frac{\frac{N_{\text{wrong}}}{T}}{\frac{N_{\text{tot}}}{T}} = \frac{R_{\text{error}}}{R_{\text{tot}}}. \tag{B1}$$

Here $R_{\text{tot}} = N_{\text{tot}}/T$ and $R_{\text{error}} = N_{\text{wrong}}/T$ are the total and the error rates, respectively, for the qubits received in the state $|0\rangle$. If R_{raw} is the overall source rate, including both $|0\rangle$ and $|1\rangle$ states, then

$$R_{\text{tot}} = \tfrac{1}{2} R_{\text{raw}}, \tag{B2}$$

since Bob sends on average an equal number of $|0\rangle$ and $|1\rangle$ states. The total rate (number and time) of qubits sent in either the $|0\rangle$ or $|1\rangle$ state (given by $f_{\text{rep}}\mu$) that were not absorbed and that managed to arrive to detectors (given by t_{link}) and were detected (given by η) is

$$R_{\text{raw}} = f_{\text{rep}} \mu t_{\text{link}} \eta. \tag{B3}$$

Here, f_{rep} is the pulse rate (the number of "attempts" to send a photon, per time) and μ is the mean number of photons per pulse. Thus, $f_{\text{rep}}\mu$ is the number of photons sent, in the unit of time. The probability that a sent photon arrives at a detector is $t_{\text{link}} \sim 10^{-\alpha L}$ (α is the absorption coefficient, and L is the transmission distance, i.e., the length of an optical cable). Finally, the detector efficiency η is the probability that a photon that arrived at a detector is actually detected. Note that $\mu \sim 0.1 \ll 1$: we ignore the low probability cases of sending two or more photons per pulse.

In general, $R_{\text{error}} = R_{\text{opt}} + R_{\text{det}}$, but here we are only interested in the error arising due to dark counts. We have

$$R_{\text{det}} = \tfrac{1}{2}\left(\tfrac{1}{2} f_{\text{rep}}\right)(1 - \mu t_{\text{link}} \eta) p_{\text{dark}} \approx \tfrac{1}{4} f_{\text{rep}} p_{\text{dark}}. \tag{B4}$$

Here, $\frac{1}{2}$ is the probability that a wrong detector will click, $(\frac{1}{2} f_{\text{rep}})$ is the rate of photons sent in the "right" state (in our case $|0\rangle$), and $(1 - \mu t_{\text{link}} \eta)$ is the probability that a photon is not detected (when dark counts are relevant). Note that the probability that a photon arrives at detectors and is efficiently detected is small, $\mu t_{\text{link}} \eta \ll 1$. Also, we neglect the less probable cases of "right" dark counts, when a photon does not arrive at detectors.

Thus, we get

$$\begin{aligned} \delta p_0(1|0) = \delta p_1(0|1) &= \frac{p_{\text{dark}}}{2\mu t_{\text{link}} \eta}, \\ \delta p_0(0|0) = \delta p_1(1|1) &= -\frac{p_{\text{dark}}}{2\mu t_{\text{link}} \eta}. \end{aligned} \tag{B5}$$

[1] G. Brassard, C. Crépeau, R. Jozsa, and D. Langlois, in *Proceedings of the 34th IEEE Foundations of Computer Science* (IEEE, Piscataway, NJ, 1993), pp. 362–371.

[2] H.-K. Lo and H. F. Chau, Phys. Rev. Lett. **78**, 3410 (1997); D. Mayers, *ibid.* **78**, 3414 (1997).

[3] A. Danan and L. Vaidman, Quantum Inf. Proc. **11**, 769 (2012).

[4] C. H. Bennett and G. Brassard, in *IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* (IEEE, Piscataway, NJ, 1984), pp. 175–179.

[5] H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999); P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000); D. Mayers, J. ACM **48**, 351 (2001); V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lutkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[6] I. B. Damgard, S. Fehr, L. Salvail, and C. Schaffner, *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2005)* (IEEE Computer Society, Washington, DC, 2005), pp. 449–458; I. B. Damgard, S. Fehr, L. Salvail, and C. Schaffner, Lect. Notes Comput. Sci. **4622**, 342 (2007); I. B. Damgard, S. Fehr, R. Renner, L. Salvail, and C. Schaffner, *ibid.* **4622**, 360 (2007); N. J. Bouman, S. Fehr, C. González-Guillén, and C. Schaffner, *ibid.* **7582**, 29 (2013).

[7] S. Wehner, C. Schaffner, and B. M. Terhal, Phys. Rev. Lett. **100**, 220502 (2008); C. Schaffner, B. Terhal, and S. Wehner, Quantum Inf. Commun. **9**, 11 (2008); R. Köenig, S. Wehner, and J. Wullschleger, IEEE Trans. Inf. Theor. **58**, 1962 (2012).

[8] N. H. Y. Ng, S. K. Joshi, C. Chen Ming, C. Kurtsiefer, and S. Wehner, Nat. Commun. **3**, 1326 (2012).

[9] A. Kent, Phys. Rev. Lett. **83**, 1447 (1999); J. Cryptolog. **18**, 313 (2005); New J. Phys. **13**, 113015 (2011); Phys. Rev. Lett. **109**, 130501 (2012); S. Croke and A. Kent, Phys. Rev. A **86**, 052309 (2012).

[10] T. Lunghi, J. Kaniewski, F. Bussieres, R. Houlmann, M. Tomamichel, A. Kent, N. Gisin, S. Wehner, and H. Zbinden, Phys. Rev. Lett. **111**, 180504 (2013); Y. Liu, Y. Cao, M. Curty, S.-K. Liao, J. Wang, K. Cui, Y.-H. Li, Z.-H. Lin, Q.-C. Sun, D.-D. Li, H.-F. Zhang, Y. Zhao, T.-Y. Chen, C.-Z. Peng, Q. Zhang, A. Cabello, and J.-W. Pan, *ibid.* **112**, 010504 (2014).

[11] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

[12] J. Bouda, P. Mateus, N. Paunković and J. Rasga, Int. J. Quantum Inf. **6**, 219 (2008).

[13] N. Paunković, J. Bouda, and P. Mateus, Phys. Rev. A **84**, 062331 (2011).

[14] H. Situ, D. Qiu, P. Mateus, and N. Paunković, arXiv:1106.3956 (quant-ph).

[15] A. Kent, Phys. Rev. Lett. **90**, 237901 (2003).

[16] A. Kent, Phys. Rev. A **68**, 012312 (2003).

[17] L. P. Lamoureux, E. Brainis, D. Amans, J. Barrett, and S. Massar, Phys. Rev. Lett. **94**, 050503 (2005).

[18] J. Barrett and S. Massar, Phys. Rev. A **70**, 052310 (2004).

[19] H. Buhrman, M. Christandl, P. Hayden, H.-K. Lo, and S. Wehner, Phys. Rev. A **78**, 022316 (2008).

[20] W. J. Munro, K. Nemoto, R. G. Beausoleil, and T. Spiller, Phys. Rev. A **71**, 033819 (2005).

[21] B. R. Johnson, M. D. Reed, A. A. Houck, D. I. Schuster, Lev S. Bishop, E. Ginossar, J. M. Gambetta, L. DiCarlo, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Nat. Phys. **6**, 663 (2010).
[22] N. A. Silva, N. J. Muga, and A. N. Pinto, IEEE J. Quantum Electron. **46**, 285 (2010).
[23] N. J. Muga, M. F. S. Ferreira, and A. N. Pinto, J. Lightwave Technol. **29**, 355 (2011).
[24] Á. J. Almeida, N. A. Silva, N. J. Muga, and A. N. Pinto, Proc. SPIE **8001**, 80013W (2011).
[25] Á. J. Almeida, N. A. Silva, P. S. André, and A. N. Pinto, Opt. Commun. **285**, 2956 (2012).
[26] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).
[27] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[28] N. J. Muga, A. N. Pinto, M. F. S. Ferreira, and J. R. Ferreira da Rocha, J. Lightwave Technol. **24**, 3932 (2006).
[29] K. Życzkowski and M. Kuś, J. Phys. A **27**, 4235 (1994).
[30] A. Bhattacharyya, Bull. Calcutta Math. Soc. **35**, 99 (1943).
[31] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[32] V. Vedral, Rev. Mod. Phys. **74**, 197 (2002).
[33] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[34] C. Gourieroux and A. Monfort, *Statistics and Econometric Models* (Cambridge University Press, Cambridge, 1995).
[35] A. Chefles, *Quantum States: Discrimination and Classical Information Transmission. A Review of Experimental Progress*, in *Quantum State Estimation*, edited by M. G. A. Paris and J. Řeháček (Springer-Verlag, Berlin, 2004), p. 467.
[36] M. Williamson and V. Vedral, J. Mod. Opt. **13**, 1989 (2003).
[37] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, MA, 2001).
[38] A. Shamir, R. L. Rivest, and L. M. Adleman, *Mental Poker* (The Mathematical Gardner, California, 1981), pp. 37–43.
[39] M. Blum, SIGACT News **15**(1), 23 (1983).
[40] S. Even, in *Advances in Cryptography: Proceedings of CRYPTO '82, Santa Barbara, CA*, edited by Allen Gersho (University of California, Santa Barbara, 1982), pp. 148–153.
[41] R. Joaquim and C. Ribeiro, Lect. Notes Comput. Sci. **7187**, 104 (2012).
[42] A. Perrig, R. Canetti, J. D. Tygar, and D. Song, *The TESLA Broadcast Authentication Protocol, RSA CryptoBytes* (RSA Security Inc., 2002), Vol. 5, pp. 1–11.
[43] C. P. Schnorr, Lect. Notes Comput. Sci. **435**, 239 (1990).
[44] J. Brown, P. Stradling, and C. H. Wittenberg, *U-Prove CTP R2 Whitepaper* (Microsoft, 2011).

Fair and optimistic quantum contract signing

N. Paunković,[1] J. Bouda,[2] and P. Mateus[1]
[1]*SQIG - Instituto de Telecomunicações, IST, TULisbon, Avenida Rovisco Pais, P-1049-001 Lisbon, Portugal*
[2]*Faculty of Informatics, Masaryk University, Botanická 68a, CZ-60200 Brno, Czech Republic*
(Received 21 September 2010; published 29 December 2011)

We present a fair and optimistic quantum-contract-signing protocol between two clients that requires no communication with the third trusted party during the exchange phase. We discuss its fairness and show that it is possible to design such a protocol for which the probability of a dishonest client to cheat becomes negligible and scales as $N^{-1/2}$, where N is the number of messages exchanged between the clients. Our protocol is not based on the exchange of *signed* messages: Its fairness is based on the laws of quantum mechanics. Thus, it is abuse free, and the clients do not have to generate new keys for each message during the exchange phase. We discuss a real-life scenario when measurement errors and qubit-state corruption due to noisy channels and imperfect quantum memories occur and argue that for a real, good-enough measurement apparatus, transmission channels, and quantum memories, our protocol would still be fair. Apart from stable quantum memories, the other segments of our protocol could be implemented by today's technology, as they require in essence the same type of apparatus as the one needed for the Bennett-Brassard 1984 (BB84) cryptographic protocol. Finally, we briefly discuss two alternative versions of the protocol, one that uses only two states [based on the Bennett 1992 (B92) protocol] and the other that uses entangled pairs, and show that it is possible to generalize our protocol to an arbitrary number of clients.

DOI: 10.1103/PhysRevA.84.062331 PACS number(s): 03.67.Dd, 03.67.Hk

I. INTRODUCTION

Contract signing [1] is an important security task with many applications, namely to the stock market and others [2]. It is a two-party protocol between Alice and Bob who share a common contract and want to exchange each other's commitments to it, thus binding to the terms of the contract. Usually, commitment is done by signing the contract on the spot.

With technology development, situations when parties involved are physically far apart become more relevant every day—distant people can communicate using ordinary or e-mail, the Internet, etc. This poses new challenges to the problem. Forcing spatially distant parties to exchange signatures opens the possibility of fraud: Bob may get the commitment from Alice (a copy of the contract with her signature on it) without committing himself, which creates an *unfair situation*. Indeed, having Alice's commitment enables Bob to appeal to a judge to bind the contract (i.e., to enforce it, declare it valid), by showing Alice's commitment to the contract (together with his signature stamped on it). On the other hand, although Alice did commit, she cannot prove it (prove that she sent her commitment to Bob) and thus cannot appeal to a judge. Moreover, she cannot prove that she did not receive Bob's commitment, so he can safely claim that he behaved honestly and sent his commitment to Alice. Note that initially Bob did not commit, but having Alice's commitment puts him in a position to *later in time* choose whether to sign the contract or not, and thus bind it or not, while Alice has no power to do either of the two.

This situation is particularly relevant in a stock market, for example, where prices of stocks may range over time, but agents must commit *beforehand* to sell or buy at a certain time in the *future*, for previously fixed prices. The unfairness allows Bob to make risky decisions in the stock market without being bound to them, unless he decides so. The problem when distant parties wish to commit to a common contract lies in the impossibility for an agent, say Alice, to prove whether she has indeed committed to it or not. Often, the contract-signing problem is said to be a variant of the so-called *return receipt* (or *certified mail*) *problem*, in which parties involved exchange mails between each other asking for proof confirming whether the other side received the message, or not.

A simple solution to this unfair situation is to have a trusted third party (usually referred to as Trent) mediating the transaction—Alice and Bob send their commitments to Trent, who then returns the receipts to the senders, and performs the message exchange *only* upon receiving both of the commitments. However, Trent's time and resources are expensive and should be avoided as much as possible. Unfortunately, it has been shown that there is no fair and viable contract-signing protocol [1,3], unless during the signature exchange phase the signing parties communicate with a common trusted agent, i.e., Trent. By *fair* protocol we mean that either both parties get each other's commitment or none gets. By *viable* protocol we mean that, if both parties behave honestly, they will both get each other's commitments.

The essence of the proof of the above impossibility result is rather simple and is related to the impossibility of establishing distributed consensus in asynchronous networks [3]. The simple assumption we need for the proof is the following: The protocol consists of a number of messages exchanged between the two parties, so that eventually, upon the termination of the protocol, both Alice and Bob acquire the signature of the other. This can be done either by sending pieces of signatures in each message, or in more sophisticated scenarios, by sending partial information needed to calculate, upon running a complex algorithm, the signature needed [4,5]. We see that if such protocol existed, it would have a final step where one message is exchanged, say, from Bob to Alice. In that case, before sending his last message, Bob would already have all the information required for him to compute Alice's signature

 ©2011 American Physical Society

of the contract (and Alice does not). Therefore, if he does not send the last message, the protocol reaches an unfair state. Note the essential importance of the asynchronicity—it is not possible for *distant* parties to arrange in advance that messages are sent *simultaneously*.

One way to come around this difficulty is to consider *optimistic* protocols that do not require communication with Trent unless something wrong comes up (some message is missing, etc.) [8]. In such protocols, the clients contact Trent regarding the given contract before the actual signing takes place. Trent notifies the request and assigns the particular contract with the clients, this way *initializing* the signing protocol. After that, the clients *exchange* messages between each other such that, if the protocol is executed correctly by both sides, both will end up with signed messages. In case something goes wrong (a message that is not according to the protocol is sent, or communication is interrupted), Trent is contacted in order to *bind* the contract.

Another work-around is to relax the fairness condition probabilistically. *Probabilistic fairness* allows one agent to have at most ε more probability of binding the contract over the other agent, at each step of the protocol. In this case, for an arbitrarily small ε, classical solutions have been found where the number of exchanged messages between the agents is minimized [6,7]. In addition to being (probabilistically) fair, in the protocol by Rabin [6] the joint probability that one agent can bind the contract, while the other cannot, is also always smaller than given ε. Moreover, the second protocol by Ben-Or *et al.* [7] satisfies an even stronger condition: The conditional probability that an agent cannot bind the contract, given that the other can, can be made arbitrarily small. Note that the two notions are not exclusive: The protocol [7] is both fair and optimistic.

In this paper, we present a (probabilistically) *fair* contract-signing protocol where *no information* with a trusted third party (Trent) is exchanged during the exchange phase. This way, it avoids possible communication bottlenecks that are otherwise inherent when involving Trent. Information exchange takes place during the initialization phase and possibly later during the (contract) binding phase (the protocol is *optimistic* [8]: Trent is rarely asked to bind the contract due to protocol fairness—cheating does not pay off). Unlike previous classical proposals, in our quantum protocol the messages exchanged between the clients (Alice and Bob) during the exchange phase do *not* have to be *signed*. Consequently, our protocol is abuse free [9]: A client has no proof that (s)he communicated with the other client, attempting to sign a *given* contract. In our protocol only two signed messages are exchanged. This is very important when one wants to achieve unconditional security. In the case of classical protocols, digital pseudosignatures [10] should be used, where the key is one-use and expensive to generate. Finally, our protocol can be used in solving some purely quantum protocols involving timely decisions between spatially distant parties; such was the case of simultaneous dense coding and teleportation [11].

In classical cryptography the contract exchange is done in the way that respective participants are learning some information (signed message, etc.) bit by bit, thus increasing their knowledge. In order to bind the contract they have to present the (complete) information to Trent. Our approach is somewhat different and is based on the laws of quantum physics.

Quantum systems obey the laws of quantum physics, which exhibit some counterintuitive features that are quite distinct from those of classical physics. The principles of quantum superposition, entanglement, and interference, to name just a few, have found numerous applications in the growing field of quantum information and computation, bringing about many advantages of quantum-based information-processing protocols over those using classical systems only. Quantum cryptography [12,13] guarantees secure communication, and devices based on its principles can be already bought on a market today; Shor's algorithm for factoring [14] is exponentially faster than any known classical counterpart; the use of entanglement can in some cases of distributed-information-processing protocols considerably decrease the amount of information exchange between distant parties needed to perform a given task, thus decreasing the (quantum) communication complexity of the problem [15] and can even eliminate any need of communication that is classically necessary for achieving a goal of common interest of separated parties [16] (a feature often referred in the literature to as *pseudotelepathy*). For an overview of the field of quantum information and computation, see for example [17].

In solving the contract-signing problem, we use quantum complementarity which as a consequence has the impossibility to unambiguously discriminate between two quantum states unless they are mutually orthogonal [18,19]. This fact is at the heart of the security of quantum cryptographic protocols. In particular, as the setups of the famous Bennett-Brassard 1984 (BB84) quantum-key-distribution protocol [12] and our contract-signing protocol have conceptual similarities, so the security of the former is closely related to the fairness of the latter. At the end, we show an alternative version of our contract-signing protocol that uses quantum entanglement, a nonlocal version of complementarity.[1] Entanglement provides the possibility of replacing the exchange of classical information with Trent during the exchange phase, by establishing the *proper* type of correlations between the distant parties, in a similar fashion as in Ekert's key distribution protocol that uses entangled states [20].

The very same mechanism of commitment to one specific choice can be used to establish e.g., a bit commitment protocol. Note that unconditionally secure bit commitment is not possible without Trent [21,22], although it is realizable using other assumptions as well.

The paper is organized as follows. In the next section we present the quantum-contract-signing protocol that requires no communication with Trent during the exchange phase and is optimistic and fair. In Sec. III we present the conditions we require for the protocol to be fair. In Sec. IV we discuss its fairness for the case of ideal measurements restricted to only two alternative single-particle observables. In Sec. V we analyze the case of general multiparticle measurements (POVMs) and argue that under the assumption of noisy

[1]Both quantum complementarity and quantum entanglement could be seen as consequences of the Heisenberg uncertainty principle and, ultimately, the superposition principle.

channels, imperfect quantum memories, and realistic detectors with error rates, it is still possible to design a fair protocol. We also discuss alternative protocols that, instead of four states, use either two nonorthogonal states or entangled pairs. Finally, we present a generalization to the case of more than two agents. In our concluding section we present a short overview of the results presented and some possible future lines of research.

II. THE PROTOCOL

In order for it to be fair, any contract-signing protocol has to force a client to make *only one* out of two possible choices—accept or reject the contract. A conceptually similar situation occurs in quantum physics, where an observer cannot simultaneously measure two complementary observables, say the position and the velocity of a quantum system. In other words, an observer is forced to choose to measure *only one* out of two possible observables and gain information about only one out of two physical properties of the system observed. For example, by measuring the exact position of a quantum particle, we are left with complete uncertainty of its velocity, and vice versa (Heisenberg uncertainty relations).

This basic feature of quantum physics is the essential ingredient of our protocol: Instead of sending to Trent the information that explicitly states the acceptance or rejection of the contract, Alice reveals her choice by measuring one of the two complementary observables, thus acquiring information about only one of the two possible features of the system given to her by Trent (and the same for Bob). Gaining information about one feature thus corresponds to the acceptance, while acquiring information about the other corresponds to rejection of the contract. This information can later be used as a proof of the client's choice.

As a client's measurement is local, no information is exchanged between a client and Trent during the exchange phase. Only later, during the possible binding phase, this (classical) information obtained by a client's measurement is confronted with Trent's (classical) information of the quantum state in which he prepared the quantum system distributed to a client, and thus used as verification of the client's choice.

To ensure the timely decisions, Trent provides Alice with the classical information of the quantum state in which Bob's quantum system is prepared, and vice versa. This way, the clients can confront each other's measurement results with the classical data provided by Trent, thus obtaining each other's commitment choices before a certain fixed moment in time. Since quantum mechanics is essentially a probabilistic theory, the clients are supplied by a number of systems, giving rise to the probabilistic fairness of the protocol.

In our protocol, we use the simplest two-dimensional quantum systems called qubits. The complementary observables could be seen as spin components (for electrons), or linear polarizations (for photons), along two mutually orthogonal axes. We will denote the two observables measured on single qubits as *the Accept* observable $\hat{A}$ and *the Reject* observable $\hat{R}$. Measuring $\hat{A}$ corresponds to the acceptance, while measuring $\hat{R}$ corresponds to the rejection of the contract. The two observables $\hat{A}$ and $\hat{R}$ are required to be mutually complementary and are given by mutually unbiased bases [18] $\mathcal{B}_A = \{|0\rangle,|1\rangle\}$ (*the Accept* basis) and $\mathcal{B}_R = \{|-\rangle,|+\rangle\}$ (*the Reject* basis), respectively, such that $|\pm\rangle = (|1\rangle \pm |0\rangle)/\sqrt{2}$. Both observables have the same eigenvalues, 0 and 1, such that

$$\begin{aligned} \hat{A} &= 1|1\rangle\langle 1| + 0|0\rangle\langle 0|, \\ \hat{R} &= 1|+\rangle\langle +| + 0|-\rangle\langle -|. \end{aligned} \tag{1}$$

During the initialization phase, Trent randomly prepares qubits, each in one of the four states $\{|0\rangle,|1\rangle,|-\rangle,|+\rangle\}$, taken from the Accept or the Reject basis. Thus, each qubit state $|\psi\rangle \in \{|0\rangle,|1\rangle,|-\rangle,|+\rangle\}$ is defined by two classical bits $C = (C_b, C_s)$, the first of which defines the basis ($C_b = 1$ if $|\psi\rangle \in \mathcal{B}_A$, while $C_b = 0$ otherwise) while the second defines the particular state from a given basis ($C_s = 1$ if $|\psi\rangle \in \{|1\rangle,|+\rangle\}$, while $C_s = 0$ otherwise). For each qubit sent to Alice in a state $|\psi\rangle^{\mathcal{A}}$, Trent sends to Bob classical bits $C^{\mathcal{A}} = C(|\psi\rangle^{\mathcal{A}})$ assigned to this state, and analogously for Bob.

During the exchange phase, which might take place much later after the initialization phase, the clients agree upon the actual contract, decide whether to commit to it or not, and exchange each other's commitments. On the system in the state $|\psi\rangle^{\mathcal{A}}$, Alice measures either $\hat{A}$ or $\hat{R}$, depending on whether she decides to accept or reject the contract, respectively, and then communicates her result $M^{\mathcal{A}}$ to Bob. Since the two observables are defined by a pair of mutually unbiased bases, the statistics of measurement results performed on a sequence of qubits each of which is randomly prepared in one of the four states $\{|0\rangle,|1\rangle,|-\rangle,|+\rangle\}$ will be dramatically different depending on which observable is measured. For instance, when Alice measures $\hat{A}$ on qubits prepared in the Accept basis, the measurement outcomes are perfectly correlated with the corresponding classical bits $C_s^{\mathcal{A}}$ given to Bob (i.e., $M^{\mathcal{A}} = C_s^{\mathcal{A}}$); when she measures $\hat{A}$ on qubits prepared in the Reject basis produces, the results obtained are completely uncorrelated with the corresponding classical bits, and analogously for $\hat{R}$.

This way, by choosing one of the two measurements performed on a sequence of qubits, Alice produces one of two mutually exclusive sets of measurement outcomes that serve as a signature of her choice (to accept or to reject the contract). By sending the results to Bob, she informs him of her decision by some fixed moment in time t_0. The same is done by Bob.

In the binding phase, *each* party is asked to confront her or his measurement results with Trent's corresponding classical bits. The perfect correlation between measurement results and the corresponding classical information for qubits prepared in the Accept or Reject basis confirms a client's Accept or Reject choice. Trent declares the contract valid, giving the signed certificates to both clients, if a client, say Alice, accepted the contract, while Bob did not reject it, or vice versa. Note that it is impossible, unless with negligible probability, to produce perfect correlations on both sets of qubits (a direct consequence of the security [13] of the BB84 protocol [12]). In this and the next two sections we restrict ourselves to the case of ideal measurements and noiseless quantum channels (from Trent to the clients). The real-case scenario is discussed in the penultimate section of the paper.

Obviously, the exchange phase as described above suffers from the same problem as any classical protocol not involving information exchange with Trent: Upon receiving Alice's

results, Bob can stop communication and safely postpone his decision to accept or reject the contract to a later moment in time. Thus, we require that the measurements and the exchange of measurement outcomes between clients happens in steps, (qu)bit by (qu)bit. Nevertheless, if we require perfect correlations between the measuring results and the corresponding classical information for *all* qubits distributed to a client, in $1/4$ of the cases an agent will be already after the first step in a position to choose to, with probability one, either bind or reject the contract.

Imagine the following situation. Alice is the first to perform a measurement and send the classical information to Bob. She is the honest party, she wants to accept the contract, and thus she measures $\hat{A}$ onto her first qubit. With probability $1/2$, her first qubit will be prepared in one of the states from the Reject basis. By measuring $\hat{A}$ onto one of the $\{|-\rangle,|+\rangle\}$ states, there is $1/2$ of the probability that her result does not match the value of $C_{s_1}^{\mathcal{A}}$, and so with probability $1/4$ Alice will not be able to achieve perfect correlations between her measurement outcomes and the corresponding classical data of the qubits prepared in states from the Reject basis. In other words, there is $1/4$ of probability that an honest side will already after the first step of the exchange phase be unable to reject the contract, even before the other side performed any measurement. In such a situation, Bob can safely stop communication and postpone his choice until a later moment in time.

Thus, we require that in order to accept/reject the contract, a client has to establish perfect correlations on $\alpha N_{A/R}$ qubits prepared in the Accept/Reject basis (with the total number of qubits $N = N_A + N_R$), where $1/2 < \alpha < 1$. We call α the *acceptance ratio* (note that for a protocol to be viable, it is necessary that $\alpha > 1/2$). In this scenario, a client is allowed to obtain $(1-\alpha)N_{A/R}$ wrong results for the states prepared in the basis of her or his choice (Accept or Reject). If α is sufficiently large, this eliminates the possibility of obtaining good-enough correlations for both groups of qubits, this way eliminating the above-mentioned unfair situation.

In the rest of this section, we present a formal description of our protocol. Our protocol is optimistic and as such it is divided into three phases: the initialization, the exchange (see Fig. 1), and the binding phase. During the exchange phase agents exchange their measurement results. If both clients are honest and perform measurements according to the protocol (measure the Accept observable), the exchange phase will end up with both clients having their probabilities to bind the contract *exponentially* (in number of qubits) close to one: The protocol is viable and optimistic (clients do not need to contact Trent as they already know the answer). If a client, say Bob, is dishonest and performs measurements other than that prescribed by the protocol (or just guesses the outcomes), he will unavoidably obtain wrong outcomes for some of the qubits from the Accept basis. If Alice detects a wrong result (Bob is *cheating*), she interrupts the exchange and proceeds to the binding phase. In a realistic case of measurement errors, Alice will have to set a threshold for the allowed number of wrong results below which she continues with the exchange. We discuss it at the end of this paper.

The initialization phase. Trent produces N pairs of qubits in states $(|\psi\rangle_m^{\mathcal{A}},|\psi\rangle_m^{\mathcal{B}})$ with the corresponding classical description $(C_m^{\mathcal{A}},C_m^{\mathcal{B}}) = [(C_{b_m}^{\mathcal{A}},C_{s_m}^{\mathcal{A}}),(C_{b_m}^{\mathcal{B}},C_{s_m}^{\mathcal{B}})]$, with $m \in \{1,\ldots,N\}$.

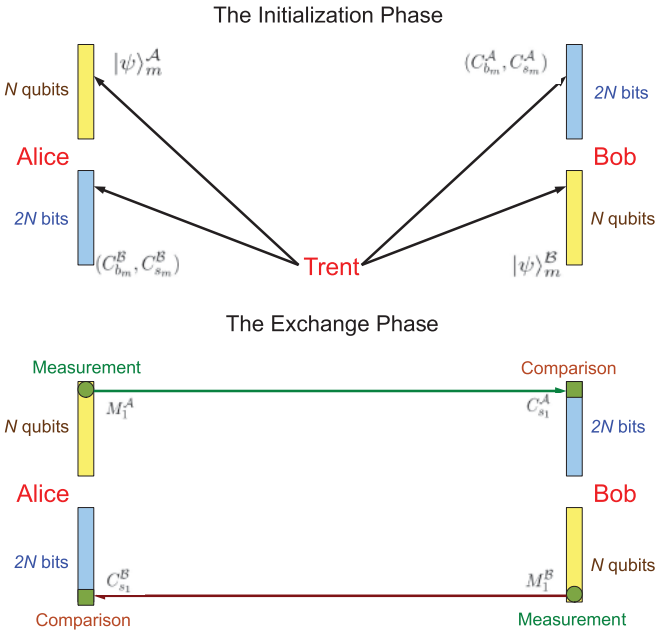


FIG. 1. (Color online) The initialization and the exchange phase.

The rule of assigning the classical data to the corresponding qubit states is the following: $C_{b_m}^{\mathcal{A}/\mathcal{B}} = 1$ if $|\psi\rangle_m^{\mathcal{A}/\mathcal{B}} \in \mathcal{B}_A$, while $C_{b_m}^{\mathcal{A}/\mathcal{B}} = 0$ otherwise; $C_{s_m}^{\mathcal{A}/\mathcal{B}} = 1$ if $|\psi\rangle_m^{\mathcal{A}/\mathcal{B}} \in \{|1\rangle,|+\rangle\}$, while $C_{s_m}^{\mathcal{A}/\mathcal{B}} = 0$ otherwise. Each qubit state is randomly chosen from the set $\{|0\rangle,|1\rangle,|-\rangle,|+\rangle\}$. Trent distributes to Alice N qubits $|\psi\rangle_m^{\mathcal{A}}$ and $2N$ classical bits $C_m^{\mathcal{B}}$, and analogously for Bob, keeping the copy of the classical data to himself. He assigns a unique identifier (number) to all data so that it can be linked in the exchange phase to a specific contract.

The exchange phase. Alice and Bob agree on a contract and exchange signed messages containing the contract, the identifier of qubits sequence they want to use, and some previously arranged moment in time t_0 giving time restriction to finish the exchange phase. (This does not bind them to the contract.) Alice and Bob perform measurements on their qubits and exchange the measurement results with each other. Without the loss of generality, we assume Alice is the first to start communication. She measures an observable of her choice ($\hat{A}$ or $\hat{R}$) on the state $|\psi\rangle_1^{\mathcal{A}}$, obtaining the result $M_1^{\mathcal{A}} \in \{0,1\}$, and sends it to Bob. Bob compares $M_1^{\mathcal{A}}$ with $C_{s_1}^{\mathcal{A}}$. If the values are different, Alice measured her qubit in the basis corresponding to $(1 + C_{b_1}^{\mathcal{A}})$ mod 2. Otherwise, the comparison is inconclusive. Next, Bob repeats the procedure described for Alice. The rest of the exchange consists in repeating the above procedure for the states $(|\psi\rangle_m^{\mathcal{A}},|\psi\rangle_m^{\mathcal{B}})$ with $m \in \{2,\ldots,N\}$. If a client, say Alice, does not obtain a result from Bob until t_0 or receives for a qubit from the Accept basis a result different from the corresponding classical data $(C_{b_m}^{\mathcal{B}} = 1 \wedge M_m^{\mathcal{B}} \neq C_{b_m}^{\mathcal{B}})$, she immediately proceeds to the binding phase.

The binding phase. At the beginning of the binding phase Trent chooses $\alpha \in (1/2,1)$ randomly and independently, according to a publicly known probability distribution $p(\alpha)$. Without the loss of generality, we assume that Alice contacts Trent to decide validity of the contract. She decides according to her preference whether she want to bind or to reject the contract. In the former case she measures all unmeasured

qubits in the Accept basis, in the latter in the Reject base. Both parties then report to Trent for each respective qubit whether they measured it in the Accept or Reject basis, and submit respective measurement outcomes. Trent verifies whether their measurement outcomes correspond to their claims. If there is a mismatch in measurements of, say, Bob, he is declared a cheater and Trent considers only Alice's measurement outcomes. Let $N_A^{\mathcal{A}}$ ($N_R^{\mathcal{A}}$) denote the number of Alice's qubits prepared in the Accept (Reject) basis, and analogously $N_A^{\mathcal{B}}$ and $N_R^{\mathcal{B}}$ for Bob. The contract is declared valid if Alice presents at least $\alpha N_A^{\mathcal{A}}$ accept results and Bob presents less than $\alpha N_R^{\mathcal{B}}$ reject results, or when Bob presents at least $\alpha N_A^{\mathcal{B}}$ accept results and Alice presents less than $\alpha N_R^{\mathcal{A}}$ reject results. In case a client, say, Bob, supplied incorrect measurement outcomes (see above), Trent declares the contract to be valid if Alice presents at least $\alpha N_A^{\mathcal{A}}$ accept results. In all other cases the contract is declared invalid.

III. FAIRNESS CONDITIONS

As noted in the previous section, our protocol is (probabilistically) viable: If both clients are honest, the probability to bind the contract is exponentially close to one. Therefore, it is also optimistic: Honest clients do not need to contact Trent in order to obtain the verdict that, with exponentially high probability, they already know. In case a client, say Bob, is not honest, i.e., is not measuring the Accept observable in every step of the protocol, we say that he is cheating. Any cheating strategy will inevitably have a nonzero probability of producing a wrong result on qubits from the Accept basis, thus allowing Alice to detect Bob's cheating and move on to the binding phase.

In case the exchange phase is terminated due to cheating detected, all we can predict is the probability that Trent declares the contract valid. This probability depends on the moment when the exchange was aborted as well as on actions of both parties (before and after the exchange was terminated). The preferences of the signing parties may change (due to commodity price changes, etc.) before it is possible to reach Trent. We say that parties are symmetric if the probability that Trent declares the contract valid is (almost) the same regardless of whether honest Alice wants to bind the contract and Bob wants to reject it, or vice versa. Note that we do not care about trivial cases when both want to reject or both to bind the contract.

This notion of symmetry is close to the weak coin tossing problem [23]: If both Alice and Bob want the same outcome (0 or 1), there is no need to guarantee unbiased coin toss. On the other hand, it is vital to assure as little bias as possible, if their preferences are contradictory.

In this section, we present our formal fairness conditions that allow for such symmetric positions of the clients. In the following, we assume Alice to be an honest and Bob a dishonest client.

The *probability to bind the contract*, P_{bind}, is a probability that during the binding phase Trent, upon receiving classical information from both Alice and Bob (i.e., their measurement results), declares the contract valid and hands in to both Alice and Bob signed (i.e., validated) copies of the contract.

The probability to bind is a function of step m, after which the communication is broken (typically, if cheating is noticed, or a connection between the clients broken). It is also a function of clients' strategies. In this section, we will consider the case of one-particle measurements only. As shown in Sec. V, the general case of multiparticle measurements, acting on at most L qubits, can be reduced to the case of one-qubit measurements, by considering the blocks of L qubits as units and then applying the same reasoning to the sequence of N/L such units. The case of global measurements on all N qubits is discussed at the end of Sec. V.

Restricting ourselves to one-qubit measurements, for a fixed parameter α, the probability to bind is a function of clients' measurements before the step m, $X^{\mathcal{A}}$ for Alice, $X^{\mathcal{B}}$ for Bob, and their behavior after the step m, $Y^{\mathcal{A}}$ for Alice, $Y^{\mathcal{B}}$ for Bob. Thus,

$$P_{\text{bind}} = P_{\text{bind}}(m,\alpha,(X^{\mathcal{A}},X^{\mathcal{B}}),(Y^{\mathcal{A}},Y^{\mathcal{B}})). \tag{2}$$

Alice is an honest client, so her strategy is $X^{\mathcal{A}} = \hat{A}^{\otimes m}$. In Sec. V we show that, when measured on qubits prepared in one of the states from the Accept and the Reject bases, general one-qubit measurements produce probability distributions equivalent to those obtained when only the Accept and the Reject observables are measured. Therefore, Bob's strategy, in which he measures the $\hat{A}$ observable $(m-\delta m)$ times, while the $\hat{R}$ observable is measured δm times (the order is irrelevant), is $X^{\mathcal{B}} = \hat{A}^{\otimes(m-\delta m)}\hat{R}^{\otimes(\delta m)}$. Thus, we can write $X^{\mathcal{B}} = X^{\mathcal{B}}(m,\delta m)$.

The only two relevant strategies after step m are those in which one client wants to bind the contract and the other does not want to bind the contract. The one who wants it bound measures the Accept observable, $Y = \hat{A}^{\otimes(N-m)}$. The other, who does not want it bound, measures the Reject observable, $Y = \hat{R}^{\otimes(N-m)}$. If Alice wants the contract bound, $Y^{\mathcal{A}} = \hat{A}^{\otimes(N-m)}$, and Bob does not, $Y^{\mathcal{B}} = \hat{R}^{\otimes(N-m)}$, we call the corresponding probability Alice's probability to bind the contract, $P_{\text{bind}}^{\mathcal{A}}(m,\alpha,X^{\mathcal{A}},X^{\mathcal{B}})$, and analogously for Bob (note that in this notation, clients' strategies after step m are not explicitly written). We will calculate clients' probabilities to bind the contract in the next section. Here, we only note that we do not discuss cases when after step m both clients want the same: If the protocol is fair for the cases when clients' wishes are opposite (a conservative assumption), it will be when they wish the same.

Averaging over all of Bob's possible strategies $X^{\mathcal{B}}$ (Alice is honest, so her strategy is known), we obtain

$$P_{\text{bind}}^{\mathcal{A}}(m,\alpha) = \sum_{X^{\mathcal{B}}} p(X^{\mathcal{B}}) P_{\text{bind}}^{\mathcal{A}}(m,\alpha,X^{\mathcal{A}},X^{\mathcal{B}}), \tag{3}$$

and analogously for $P_{\text{bind}}^{\mathcal{B}}(m,\alpha)$. Here, $p(X^{\mathcal{B}})$ is the probability that Bob chooses the particular strategy $X^{\mathcal{B}} = X^{\mathcal{B}}(m,\delta m)$, and is given by the probability $p_w(\delta m) = 1-(3/4)^{\delta m}$ to obtain a wrong result, when measuring the Reject observable δm times (obviously, wrong results are, in the case of ideal measurements with no errors, possible only on qubits from the Accept basis).

For our protocol to be fair, we require that at each step m of the exchange phase, the difference between the agents'

(Alice and Bob) probabilities to bind the contract can be made arbitrarily small: For any given ε,

$$\left|P^{\mathcal{B}}_{\text{bind}}(m,\alpha) - P^{\mathcal{A}}_{\text{bind}}(m,\alpha)\right| < \varepsilon. \tag{4}$$

In order to make our protocol even more symmetric, we introduce *the probability to cheat* of a dishonest client (Bob). It is the product of Bob's probability to bind and the probability that Alice will not bind the contract:

$$P^{\mathcal{B}}_{ch}(m,\alpha,X^{\mathcal{A}},X^{\mathcal{B}}) = P^{\mathcal{B}}_{\text{bind}}(m,\alpha,X^{\mathcal{A}},X^{\mathcal{B}}) \times \left[1 - P^{\mathcal{A}}_{\text{bind}}(m,\alpha,X^{\mathcal{A}},X^{\mathcal{B}})\right]. \tag{5}$$

After averaging over Bob's strategies $X^{\mathcal{B}}(m)$, the probability to cheat of a dishonest client Bob is

$$P^{\mathcal{B}}_{ch}(m,\alpha) = \sum_{X^{\mathcal{B}}} p(X^{\mathcal{B}}) P^{\mathcal{B}}_{ch}(m,\alpha,X^{\mathcal{A}},X^{\mathcal{B}}). \tag{6}$$

Our second fairness requirement is that a dishonest client's (Bob's) probability to cheat $P^{\mathcal{B}}_{ch}(m,\alpha)$ is also negligible.[2] Note that in our protocol the binding phase requires *both* clients to confront their measurement results, *both* obtaining the *same* verdict by Trent at the end. Therefore, although the probability to cheat is the product of two probabilities, the probability that Bob can bind the contract and the probability that Alice cannot, it is not itself a probability of an event (Bob's probability to bind is obtained under the assumption that after step m he measures $\hat{A}$ and Alice $\hat{R}$; to obtain Alice's probability to bind, we assume that after step m she is the one to measure $\hat{A}$ while Bob measures $\hat{R}$). Yet, it can serve as a measure of a protocol's fairness as it quantifies an agent's freedom to choose between binding and refusing the contract later in time. We discuss this in more detail at the end of this section. At the end of the paper, we show that it is straightforward to design a protocol in which it is not necessary that both clients are (simultaneously) present during the binding phase. In this scenario, the corresponding probability to cheat becomes the joint probability that one client can bind the contract, while the other cannot.

Unfortunately, although the first fairness criterion is satisfied for the above protocol, its probability to cheat can be as high as 1/4: If both clients measure the Accept observable, both clients' probabilities to bind are monotonic functions, such that for big enough N there exists m_s for which $P^{\mathcal{A}}_{\text{bind}}(m_s,\alpha) \approx P^{\mathcal{B}}_{\text{bind}}(m_s,\alpha) \approx 1/2$. But if α, determined by Trent by sampling a random variable described by a publicly known probability distribution $p(\alpha)$, is itself unknown to the clients, and determined only later during the binding phase, the expected probability to cheat $\bar{P}_{ch}(m)$ becomes negligible, for big enough N. Thus, our second fairness condition finally reads as follows: For any given ε,

$$\bar{P}^{\mathcal{B}}_{ch}(m) = \int p(\alpha) P^{\mathcal{B}}_{\text{bind}}(m,\alpha)\left[1 - P^{\mathcal{A}}_{\text{bind}}(m,\alpha)\right] d\alpha < \varepsilon. \tag{7}$$

The coefficient α is sampled randomly by Trent to achieve stronger security requirements. This assures the symmetric position of honest and cheating participant even before Trent is contacted during the binding phase: If agents are temporarily unable to contact Trent, cheaters should not profit from this in a significant way.

A client, say Bob, may be willing to take the risk and stop the protocol prematurely during the exchange phase, provided such a situation can assure him some reasonable position. Consider a contract where Alice buys orange juice from Bob for X units per liter. According to the market expectation, with probability p the price should increase and with probability $(1-p)$ decrease. When the price goes up to $X' > X$, Alice wants to enforce the contract, since otherwise she would buy juice for a higher price. Bob wants the contract to be canceled to sell the juice for a higher price. If the price drops, the situation is symmetric.

Bob may be willing to take the risk parameterized by δ in the following sense. The joint probability that the price drops and he will be able to enforce the contract is at least δ as well as the joint probability that the price increases and Alice will not be able to enforce the contract. The latter gives him protection from financial losses, while the former allows him to spare some money. This is formalized as

$$(\exists\, 0 \leqslant p \leqslant 1\,)\left[p\left(1 - P^{\mathcal{A}}_{\text{bind}}\right) \geqslant \delta \;\wedge\; (1-p)P^{\mathcal{B}}_{\text{bind}} \geqslant \delta\right]. \tag{8}$$

Thus, to prevent reasonability of Bob's cheating we require that

$$(\forall\, 0 \leqslant p \leqslant 1)\left[p\left(1 - P^{\mathcal{A}}_{\text{bind}}\right) \leqslant \delta \;\vee\; (1-p)P^{\mathcal{B}}_{\text{bind}} \leqslant \delta\right]. \tag{9}$$

Let us denote $Y \stackrel{\text{def}}{=} P^{\mathcal{B}}_{\text{bind}}(m,\alpha)[1 - P^{\mathcal{A}}_{\text{bind}}(m,\alpha)]$ the random variable parameterized by α. The expected value

$$E(Y) = \int p(\alpha) P^{\mathcal{B}}_{\text{bind}}(m,\alpha)\left[1 - P^{\mathcal{A}}_{\text{bind}}(m,\alpha)\right] d\alpha \tag{10}$$

is nothing but the *expected probability to cheat* $\bar{P}_{ch}(m)$ [note that due to $P^{\mathcal{B}}_{\text{bind}}(m,\alpha) \approx P^{\mathcal{A}}_{\text{bind}}(m,\alpha)$, we have $\bar{P}_{ch}(m) \equiv \bar{P}^{\mathcal{B}}_{ch}(m) \approx \bar{P}^{\mathcal{A}}_{ch}(m)$]. Using the above fairness criterion $\bar{P}_{ch}(m) < \varepsilon$, Chebyshev inequality, and putting $\delta^3 = \varepsilon$, we obtain $\text{Prob}_{\alpha}[Y < \delta + \delta^3] \geqslant 1 - \delta$. Thus, the probability δ can be made arbitrarily small with arbitrarily high probability.

IV. FAIRNESS OF THE PROTOCOL: IDEAL CASE

In the following, we show that our protocol is *fair*: Both $|P^{\mathcal{B}}_{\text{bind}}(m,\alpha) - P^{\mathcal{A}}_{\text{bind}}(m,\alpha)|$ and $\bar{P}_{ch}(m)$ could be made arbitrarily low. In this section, we assume that only $\hat{A}$ or $\hat{R}$ are measured, and that no measurement errors or qubit-state corruption occur. In the next section, we discuss general (one or multiqubit) observables and a real-life scenario of imperfect measurements and noisy channels.

If Bob is cheating during the exchange phase, he will be detected after a small number of steps, with the probability growing exponentially in the number of qubits measured by Alice. Assume Bob's cheating is detected after Alice measured m qubits. Alice terminated the exchange phase and participants proceed with the binding phase, which can be delayed (Trent is off-line). Meanwhile, participants are allowed to change their preferences and we would like to examine the symmetry of their position. We are interested only in the situation when

[2]Obviously, if the second fairness requirement is satisfied, the first if satisfied as well: Having negligible probability to cheat is a stronger condition than the symmetry between the client's probabilities to bind the contract.

Alice wants to bind the protocol and Bob wants to reject, and vice versa.

In the former case Alice tries to do her best to bind the contract. This means she measures all unmeasured qubits in the Accept basis and sends her results to Trent. Bob does his best to invalidate the contract, measuring his qubits in the Reject basis. Note that possible lying about the measurement basis on the respective qubit is detected with the probability growing exponentially in the number of wrongfully reported measurement results, so the number of measurements Bob can lie about is well limited.

If the cheating was detected and the exchange phase was terminated after m steps, Alice's probability to bind the contract is, for a given α, given by the following expression:

$$P^{\mathcal{A}}_{\text{bind}}(m,\alpha,X^{\mathcal{A}},X^{\mathcal{B}}) = P^{\mathcal{A}}_{A}(m,\alpha,X^{\mathcal{A}})\big[1 - P^{\mathcal{B}}_{R}(m,\alpha,X^{\mathcal{B}})\big], \tag{11}$$

and analogously for Bob. The probability to accept the contract, P_A, is the probability that a client who wants the contract bound, thus after step m measuring the Accept observable, will pass Trent's test during the binding phase. The probability to reject the contract, P_R, is the probability that a client who does not want the contract bound, thus after step m measuring the Reject observable, will invalidate the contract during the binding phase. Note that under this convention it is redundant to specify clients' strategies Y after the step m. Also, clients' probabilities to accept and reject the contract depend only upon their own strategies before the step m, and not those of the opponent.

Alice's probability to accept the contract is $P^{\mathcal{A}}_{A}(m,\alpha) = 1$: She is an honest client and thus is measuring $\hat{A}$ in every step until the interruption; she wants the contract bound, so continues to measure the Accept observable. Her probability to reject the contract can be written in a simplified form as $P^{\mathcal{A}}_{R}(m,\alpha)$, bearing in mind that in this case $X^{\mathcal{A}}(m) = \hat{A}^{\otimes m}$ and $Y^{\mathcal{A}}(m) = \hat{R}^{\otimes(N-m)}$.

Out of m steps, a dishonest Bob measures the Accept observable $(m - \delta m)$ times, and the remaining δm times the Reject observable. Therefore $X^{\mathcal{B}}(m,\delta m) = \hat{A}^{\otimes(m-\delta m)}\hat{R}^{\otimes(\delta m)}$. When rejecting the contract, his strategy is $Y^{\mathcal{B}}(m) = \hat{R}^{\otimes(N-m)}$, and we see that Bob's probability to reject is the same as Alice's probability to reject; had the interruption occurred δm steps before, $P^{\mathcal{B}}_{R}(m,\alpha,\delta m) = P^{\mathcal{A}}_{R}(m - \delta m,\alpha)$. When measuring the Reject observable, the probability to obtain wrong results on states from the Accept basis, and thus being detected cheating, is exponentially fast approaching to one, $p_w(\delta m) = 1 - (3/4)^{\delta m}$. Therefore, for $1/2 < \alpha < 1$ and large enough N, $\delta m \ll (1-\alpha)N$. The expected δm is small and since for large enough N all the probabilities are slow functions of m, we have $P^{\mathcal{B}}_{A}(m,\alpha,\delta m) \approx P^{\mathcal{A}}_{A}(m,\alpha) = 1$, while $P^{\mathcal{A}}_{R}(m,\alpha) \approx P^{\mathcal{A}}_{R}(m - \delta m,\alpha) = P^{\mathcal{B}}_{R}(m,\alpha)$. Therefore, the first fairness condition $P^{\mathcal{A}}_{\text{bind}}(m,\alpha) \approx P^{\mathcal{B}}_{\text{bind}}(m,\alpha)$ is satisfied.

To show that the second fairness condition is also satisfied, we first note that, due to $P^{\mathcal{A}}_{R}(m,\alpha) \approx P^{\mathcal{B}}_{R}(m,\alpha)$ and $P^{\mathcal{B}}_{A}(m,\alpha) \approx P^{\mathcal{A}}_{A}(m,\alpha) = 1$, the two probabilities to cheat are almost the same and can be, for a given α, written in terms of single expected probability to reject the contract:

$$P_{ch}(m;\alpha) = P_R(m;\alpha)[1 - P_R(m;\alpha)]. \tag{12}$$

The expected probability to reject, for a given α, is

$$P_R(m;\alpha) = \sum_{N_R=0}^{N} q(N_R)P_R(m;\alpha,N_R), \tag{13}$$

where $P_R(m;\alpha,N_R)$ is the probability to (be able to) reject the contract [obtain less than $(1-\alpha)N_R$ wrong results on qubits from the Reject basis, measuring the Accept observable on the first m qubits], for a given acceptance ratio α and the number of qubits prepared in the Reject basis N_R, keeping for simplicity the N dependence implicit, and $q(N_R) = 2^{-N}\binom{N}{N_R}$ is the probability to have exactly N_R states from the Reject basis. Note that $P_R(m;\alpha,N_R)$ is calculated under the assumption that the Accept observable is measured during the first m steps, while after that the Reject observable is measured for the remaining $(N-m)$ steps.

For $m < (1-\alpha)N_R$ there is always a chance to reject the contract; thus $P_R(m;\alpha,N_R) = 1$. Otherwise,

$$P_R(m;\alpha,N_R) = \sum_{n=n'}^{m'} P(n \text{ in R};m,N_R)P_R(n \text{ in R};\alpha,N_R). \tag{14}$$

Here, the probability that exactly n out of the first m qubit states are from the Reject basis is given by[3]

$$P(n \text{ in R};m,N_R) = \binom{m}{n}\binom{N-m}{N_R-n}\binom{N}{N_R}^{-1}, \tag{15}$$

while the probability of being able to reject the contract if n qubits are from the Reject basis is given by[4]

$$P_R(n \text{ in R};\alpha,N_R) = 2^{-n}\sum_{i=0}^{T}\binom{n}{i}, \tag{16}$$

where $T = (1-\alpha)N_R - 1$ if $n \geqslant (1-\alpha)N_R$ and $T = n$ otherwise. Due to the constraint of having exactly $N_{A/R}$ qubits from the Accept/Reject basis, the range of the summation for n in Eq. (14) is given by $n' = 0$ for $m \leqslant N_A$ while $n' = m - N_A$ otherwise, and $m' = m$ for $m \leqslant N_R$ while $m' = N_R$ otherwise.

Finally, the expected probability to cheat, with respect to a given probability distribution $p(\alpha)$ on the segment I_α, is

$$\bar{P}_{ch}(m) = \int_{I_\alpha} p(\alpha)P_{ch}(m;\alpha)d\alpha. \tag{17}$$

[3]Since each sequence of N qubit states, prepared such that exactly N_R are from the Reject basis, is equally probable, the probability $P(n \text{ in R};m)$ is given by the ratio of the number of sequences whose n out of N_R qubits prepared in the Reject basis are among the first m qubits, while the other $(N_R - n)$ are among the rest of $(N-m)$ qubits, $\binom{m}{n}\binom{N-m}{N_R-n}$, and the total number of allowed sequences is $\binom{N}{N_R}$.

[4]When measuring the Accept observable $\hat{A}$ on a state from the Reject basis each of the two possible results are equally probable. Therefore, each sequence of n such measurement results has the probability 2^{-n}. If exactly i out of n individual-qubit results do not match those given by Trent to the other client, and if $i < (1-\alpha)N$, then it is still possible to reject the contract. There are $\binom{n}{i}$ of such measurement results, for given n.

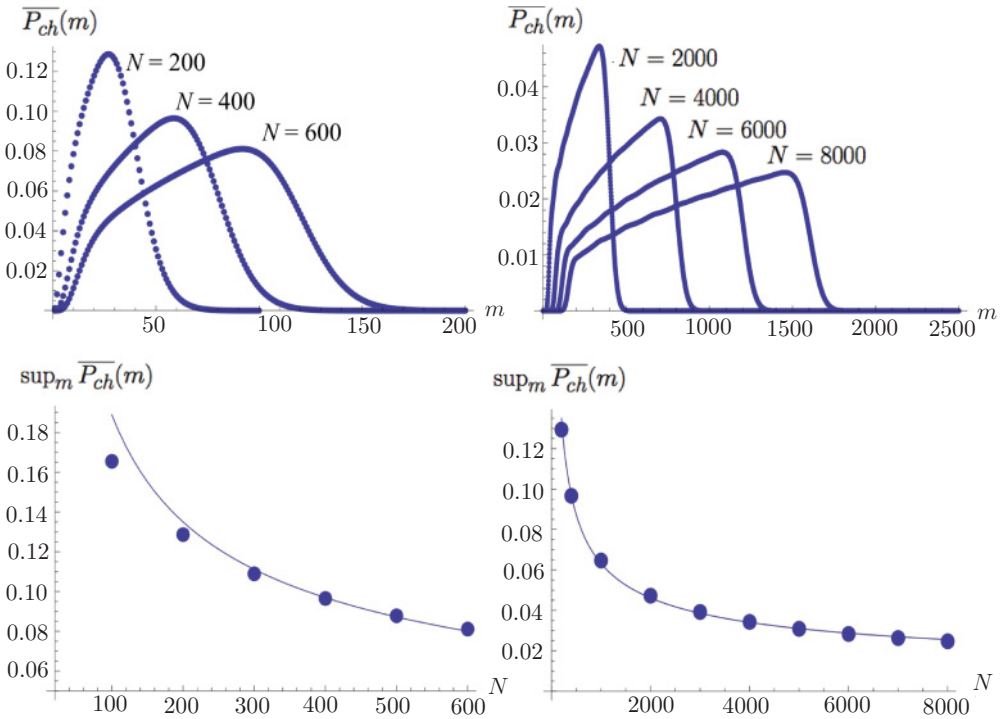


FIG. 2. (Color online) The expected probability to cheat $\bar{P}_{ch}(m)$ (upper row) and the maximal expected probability to cheat $\sup_m \bar{P}_{ch}(m)$ (lower row) for the uniform $p(\alpha)$ on $I_\alpha = [0.9,0.99]$. The plots from the left column represent results for our protocol, while the right ones are for the restricted "typical" case of $N_A = N_R$. Note the scaling behavior $\sup_m \bar{P}_{ch}(m) \propto N^{-1/2}$.

Using the simplest uniform probability $p(\alpha) = 1/I_\alpha$ on the segment $I_\alpha = [0.9,0.99]$, we numerically evaluated the expected probability to cheat $\bar{P}_{ch}(m)$ for up to $N = 600$, while for the "typical" case of $N_A = N_R$ we managed to evaluate it up to $N = 8000$; see Fig. 2. We see that the *fairness condition* $\sup_m \bar{P}_{ch}(m) \ll 1$ is satisfied [for $N = 600$ we got $\sup_m \bar{P}_{ch}(m) = \bar{P}_{ch}(92) = 0.0811$, while for $N_A = N_R$ we have $\sup_m \bar{P}_{ch}(m) = \bar{P}_{ch}(1455) = 0.0247$ for $N = 8000$), which is explicitly shown in Fig. 2. Moreover, the numerical fit gives $\sup_m \bar{P}_{ch}(m) \propto N^{-1/2}$ behavior, giving us the scaling of the complexity of the protocol, with respect to the number N of exchanged messages between Alice and Bob.

V. FAIRNESS OF THE PROTOCOL: GENERAL MEASUREMENTS AND NOISE

First, we consider only one-qubit orthogonal measurements. Since Alice is an honest client, $X^{\mathcal{A}} = \hat{A}^{\otimes m}$, we have that $P_A^{\mathcal{A}}(m,\alpha,X^{\mathcal{A}}) = 1$ and $P_R^{\mathcal{A}}(m,\alpha,X^{\mathcal{A}}) = P_R(m,\alpha)$.

Bob is a dishonest client and his strategy $X^{\mathcal{B}}$ consists of measuring $(m-k)$ times the Accept observable $\hat{A}$ and k times observable $\hat{K} = 0|m\rangle\langle m| + 1|m^\perp\rangle\langle m^\perp|$, where

$$|m\rangle = \cos\frac{\theta}{2}|0\rangle + e^{\varphi}\sin\frac{\theta}{2}|1\rangle = \cos\frac{\theta'}{2}|-\rangle + e^{\varphi'}\sin\frac{\theta'}{2}|+\rangle. \tag{18}$$

Let $m = m_a + m_r$, where m_a is the number of the Accept and m_r the number of the Reject qubits among the first m qubits, and let k_a be the number of measurements of $\hat{K}$ on m_a qubits from the Accept basis, and analogously for k_r. Bob's measurements of $\hat{K}$ are equivalent to[5] (for simplicity, we omit writing the α and N_R dependencies)

(i) $q_a k_a$ measurements of $\hat{A}$ and $\delta m_a = [1-q_a]k_a$ measurements of $\hat{R}$ on qubits from the Accept basis, where $q_a = \cos\theta$. Thus, the probability to notice cheating after k_a measurements of $\hat{K}$ is $\tilde{p}_w(\delta m_a) = \tilde{p}_w([1-q_a]k_a) = 1 - (1/2)^{[1-q_a]k_a}$ and Bob's probability to accept the contract is $P_A^{\mathcal{B}}(m,\alpha,X^{\mathcal{B}}) \equiv P_A(m,\alpha;m_a,\delta m_a)$;

(ii) $q_r k_r$ measurements of $\hat{A}$ and $\delta m_r = [1-q_r]k_r$ measurements of $\hat{R}$ on qubits from the Reject basis, where $q_r = \cos\theta'$. Thus, Bob's measurements are equivalent to $(m_r -$

[5]First, we consider measurements of $\hat{K}$ on qubits whose states are from the Accept basis. Let $p_k(0|0)$ be the conditional probability to obtain result 0 when measuring an observable $\hat{K}$ on the state $|0\rangle$, and analogously for other conditional probabilities. Notice that, when measuring an arbitrary one-qubit orthogonal observable $\hat{K}$, one has $p(0|0) = p(1|1) = s_a$: The probabilities to obtain the right result on the states from the Accept basis are the same. The two probabilities are also equal when one measures $\hat{A}$ with frequency (i.e., probability) q_a and $\hat{R}$ with frequency $1-q_a$. The probability to obtain the right result on the states from the Accept basis is then $S_a = q_a + (1-q_a)/2 = (1+q_a)/2$: When measuring $\hat{A}$ (with probability q_a) one obtains the right result with probability 1, while when measuring $\hat{R}$ (with probability $1-q_a$) one obtains the right result with probability $1/2$. In order for those two measurements to be equivalent, it is necessary that $s_a = S_a = (1+q_a)/2$. In the case of the orthogonal observable $\hat{K}$, we have $s_a = \cos^2(\theta/2)$, and thus $q_a = \cos\theta$. The same argument applies for the case of measurements performed on the states from the reject basis.

δm_r) measurements of $\hat{A}$ on qubits from the Reject basis and his probability to reject is $P_R^{\mathcal{B}}(m,\alpha,X^{\mathcal{B}}) \leqslant P_R(m-\delta m_r,\alpha)$.

Since $\tilde{p}_w(\delta m_a) = 1-(1/2)^{[1-q_a]k_a}$, either k_a is small, or $q_a \approx 1$. In the case of $q_a \approx 1$, we have that $q_r \approx 0$: The observable $\hat{K}$ is close to $\hat{A}$ and Bob's strategy is close to that of an honest client. If k_a is small, we have that δm_a is also small and Bob's probability to accept the contract is close to one, $P_A^{\mathcal{B}}(m,\alpha,X^{\mathcal{B}}) = P_A(m,\alpha;m_a,\delta m_a) \approx 1$. Moreover, since for typical cases $k_a \approx k_r$, we have that typically Bob's probability to reject the contract is $P_R^{\mathcal{B}}(m,\alpha,X^{\mathcal{B}}) \approx P_R(m,\alpha)$. Therefore, the corrected probability to cheat, averaged over all possible distributions of states (from the Accept and the Reject bases) and all possible strategies of the clients, will not be considerably altered and the protocol would still be fair, even if an arbitrary number of observables $\hat{K}_i$ is allowed. Note that, according to the above and the analysis from the previous section, the complexity of the protocol, for the case of general one-qubit orthogonal measurements, scales as $N \propto \varepsilon^{-2}$, where N is the number of exchanged messages and ε a given upper bound for the probability to cheat.

The above argument could be generalized for joint L-qubit measurements, if $L \propto N^t$ and $t < 1$: For every joint observable (or general POVM) $\hat{O}_L \neq \otimes_{i=1}^{L}\hat{A}_i$ there is a nonzero probability q_L that at least one wrong result will be obtained on the accept qubits, which scales as q_L^k, k being the number of $\hat{O}_L$ measurements (note that for ideal measurements, Alice will notice cheating as soon as she receives the *first* wrong result from Bob). Thus, in the case of performing joint measurements on at most L qubits, for $N/L \gg 1$, the protocol would still be fair.

In the case of $L \propto N$ the fairness of our protocol could be seen as a consequence of the unconditional security of the BB84 protocol in the presence of noisy channels and imperfect sources and detectors [13]: If we interpret $(1-\alpha)$ as the error rate due to the noise (see the paragraph below), then in order to successfully cheat in our contract-signing protocol, an agent would have to be correct on *both* αN_A qubits from the Accept basis and on αN_R qubits from the Reject bases, which in turn contradicts the unconditional security of the BB84 protocol. Namely, in order to pass the test by Alice and Bob and thus learn the secret key, Eve has to know the states of qubits from both mutually unbiased bases. In the presence of errors, she needs to be correct only on the α fraction of qubits. Since quantum cryptography is unconditionally secure, that is not possible, unless with exponentially small probability. Thus, our quantum-contract-signing protocol is fair. Note that although the errors due to imperfect technology are inevitable, it is possible, at least in principle, to make them arbitrarily low. Therefore, if for some error rate $(1-\alpha)$ it is not possible, unless with negligible probability, to have the right results on $\alpha N_{A/R}$ qubits from both the Accept and the Reject bases, then for some better equipment the fairness condition would be satisfied.

In the case of measurement errors, imperfect quantum memories, and noisy channels, one must introduce the error tolerance $\eta = M_w/M$, where $M_w = \langle m_w \rangle \equiv \eta M$ is the expected number of wrong results obtained in measuring an observable on M qubits prepared in states from the observable's eigenbasis. Coefficient η gives the ratio of unavoidably produced wrong results: To detect cheating would then mean to obtain more than expected, according to η, wrong results. For $\eta < (1-\alpha)$ and big enough N, our protocol would therefore still be fair.

Our protocol can be modified to use only two nonorthogonal pure states, just as the Bennett 1992 (B92) cryptographic protocol [24] is modified after the BB84. Obviously, the probabilities determining the features and complexity of the protocol would be quantitatively different, but the protocol would clearly still be fair.

We note that it is straightforward to design a protocol in which a single client can contact Trent to obtain a signed contract. In this case, Trent sends to a client, say, Alice, classical information about only half of, randomly chosen, Bob's qubit states. This information is used by Alice to check Bob's measurements: whether he is measuring the Accept observable, or not. If so, he is measuring the Accept observable on the rest of his qubits as well (he does not know for which qubits Alice has the classical information and for which not). Thus, the results Bob provided her for the rest of his qubits are used by Trent to verify Alice's data: Instead of Bob appearing with his results, Alice provides those given to her by Bob. Note that in this case, unlike before, the corresponding average probability to cheat is the probability of a real event: the joint probability that an agent cannot bind the contract, while the other can.

This version of the protocol is also suited for the use of entangled pairs, instead of qubits in definitive pure states. In this case, Trent produces $4N$ pairs of maximally entangled qubits in the singlet state $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. Half of each pair of the first $2N$ entangled pairs are distributed to Alice, while the other half is kept with Trent. Each half of the other $2N$ pairs is sent to Bob. Then, Trent randomly measures either of the two observables (Accept or Reject) on each of $4N$ qubits left with him. Random N results obtained from measuring the first $2N$ qubits are sent to Bob (together with the classical information of their positions within the first $2N$ pairs), and analogously with the other $2N$ results (N of which are sent to Alice). We see that this situation is equivalent to the above, when instead of entangled, Trent sends qubits in pure states.

Finally, we note that it is simple to generalize our protocol for the case of an arbitrary number of parties. In the case of n parties, classical information about the preparation of each of N qubits sent to, say, the first client, is divided into $N/(n-1)$ sets S_i, with $i = 2,\ldots,n$, such that each set S_i is sent to the ith client. The protocol is executed in N rounds, such that in the ith step of each round, the ith client is performing a single-qubit measurement and publicly announces the result to all of the remaining $(n-1)$ clients. To declare a contract valid, binding all of the n clients, the requirements analogous to the ones for the two-party protocol must be satisfied between each of the $\binom{n}{2}$ pairs of clients. This, as well as the above-mentioned generalizations of our protocol, will be discussed in more detail elsewhere.

VI. CONCLUSIONS

We have presented a fair and optimistic quantum protocol for signing contracts that does not require the exchange of information with the trusted party during the exchange phase.

Unlike the classical proposals, its fairness is based on the laws of physics rather than on sending signed messages, which are only computationally secure. Thus, no keys are generated during the exchange phase and the protocol is abuse free. For single-qubit orthogonal measurements, the complexity of the protocol scales as $N \propto \varepsilon^{-2}$, where N is the number of exchanged messages and ε the given threshold for the probability to cheat. Analogously to the previous proposal of a quantum-contract-signing protocol [25], the present one could also be performed using either entangled pairs or two instead of four pure states, but it does not require tamper-proof devices nor is it based on the effects of decoherence (which are necessary for achieving [25]). Also, it is simple to generalize it to involve many clients. Our protocol can easily be altered such that only one client is enough to present her or his results in order to bind the contract and obtain the document declaring the contract valid. In this case, the probability to cheat becomes a joint probability that a client can bind a contract, while the other cannot. The fairness condition of a negligible probability to cheat is in this case slightly weaker than the one of [7]. Also, in our quantum protocol the trusted party is slightly more involved than in the classical protocol presented in [6]. Nevertheless, being based on the laws of physics, our protocol has other security advantages over the classical counterparts. Moreover, it can be used for designing other quantum-information and -security protocols. Indeed, our approach was used to design the protocols of simultaneous dense coding and teleportation between spatially distant clients [11]. Finally, apart from stable quantum memories, our protocol could be easily performed with the current technology used in quantum cryptography.

The possible future lines of research include a detailed study of the effects of noise and multiparticle measurements on the fairness of the protocol and its complexity. Today's quantum technology is still in its infancy and experimental realizations and technological applications of quantum-information protocols are relatively rare and not always as reliable as classical alternatives. Even the security of quantum cryptography, which is proven to be unconditionally secure under the effects of noise, is still very much dependent on the technology used. Nevertheless, the development of quantum technology has already proven to be remarkable and beyond many of early expectations, and will for sure continue to develop. Another possible line of future research could be a further study of alternative versions of quantum-contract-signing protocols with possibly better properties and designing novel quantum-information, -security, and -computation protocols involving spatially distant clients that require timely decisions.

Note added. Recently, we became aware of a related paper on a practical quantum-bit commitment protocol published online [26].

ACKNOWLEDGMENTS

N.P. and P.M. thank the SQIG project at IT, funded by the FCT and EU FEDER projects QSec PTDC/EIA/67661/2006, QuantPrivTel PTDC/EEA-TEL/103402/2008, and FCT PEst-OE/EEI/LA0008/2011 and the IT project QuantTel, as well as Network of Excellence, Euro-NF. J.B. acknowledges the support of the Czech Science Foundation GAČR projects P202/12/1142 and P202/12/G061. The authors acknowledge discussions with V. Božin and Č. Brukner.

[1] S. Even and Y. Yacobi, Technical Report 175, Technion, Israel, 1980 (unpublished).

[2] N. Asokan, V. Shoup, and M. Waidner, IEEE J. Sel. Areas Commun. **18** (4), 593 (2000).

[3] M. J. Fischer, N. A. Lynch, and M. Paterson, J. ACM **32**, 374 (1985).

[4] S. Even, Technical Report 231, Computer Science Department, Technion, Israel, 1982 (unpublished).

[5] S. Even, O. Goldreich, and A. Lampel, Commun. ACM **28**, 637 (1985).

[6] M. Rabin, J. Comput. Syst. Sci. **27**, 256 (1983).

[7] M. Ben-Or, O. Goldreich, S. Micali, and R. L. Rivest, IEEE Trans. Inf. Theory **36**, 40 (1990).

[8] N. Asokan, M. Schunter, and M. Waidner, in *Proceedings of the 4th ACM Conference on Computer and Communications Security* (ACM, New York, 1997), p. 7.

[9] J. Garay, M. Jakobsson, and P. MacKenzie, in *Advances in Cryptology, CRYPTO '99: 19th Annual International Cryptology Conference*, edited by M. Wiener, Lecture Notes in Computer Science, Vol. 1666 (Springer, Berlin, 1999), pp. 449-466.

[10] D. Chaum and S. Roijakkers, in *Advances in Cryptology, CRYPTO '90*, edited by A. J. Menezes and S. A. Vanstone, Lecture Notes in Computer Science, Vol. 537 (Springer, Berlin, 1991), p. 206; G. Wang, IEEE Trans. Inf. Forensics Security **5**, 158 (2010).

[11] H. Situ, D. Qiu, P. Mateus, and N. Paunković, e-print arXiv:1106.3956v1.

[12] C. H. Bennett and G. Brassard, *IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), p. 175.

[13] H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999); P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000); D. Mayers, J. ACM **48**, 351 (2001); V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lutkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[14] P. W. Shor, in *Proceedings of the 35nd Annual Symposium on the Foundations of Computer Science* (IEEE Computer Society Press, Los Alamitos, CA, 1994), p. 124.

[15] Č. Brukner, M. Zukowski, J.-W. Pan, and A. Zeilinger, Phys. Rev. Lett. **92**, 127901 (2004); H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, e-print arXiv:0907.3584.

[16] G. Brassard, A. Broadbent, and A. Tapp, in *Proceedings of the 8th International Workshop on Algorithms and Data Structures*, Lecture Notes in Computer Science, Vol. 2748 (Springer, Berlin, 2003), p. 1; Č. Brukner, N. Paunković, T. Rudolph, and V. Vedral, Int. J. Quant. Info. **4**, 365 (2006).

[17] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[18] I. D. Ivanović, J. Phys. A **14**, 3241 (1981).

[19] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[20] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[21] H.-K. Lo and H. F. Chau, Phys. Rev. Lett. **78**, 3410 (1997).
[22] D. Mayers, in *PhysComp '96: Proceedings of the Fourth Workshop on Physics and Computation*, edited by T. Toffoli and M. Biafore (North-Holland, Amsterdam, 1998).
[23] C. Mochon, e-print arXiv:quant-ph/0403193.
[24] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).
[25] J. Bouda, P. Mateus, N. Paunković, and J. Rasga, Int. J. Quant. Inf. **6**, 219 (2008).
[26] A. Danan and L. Vaidman, Quantum Inf. Process. (2011), doi:10.1007/s11128-011-0284-4.

Fidelity spectrum and phase transitions of quantum systems

P. D. Sacramento,[1] N. Paunković,[2] and V. R. Vieira[1]

[1]*Departamento de Física and CFIF, Instituto Superior Técnico, TU Lisbon, Avenida Rovisco Pais, P-1049-001 Lisboa, Portugal*

[2]*SQIG-Instituto de Telecomunicações, IST, TU Lisbon, Avenida Rovisco Pais, P-1049-001 Lisboa, Portugal*

(Received 29 July 2011; published 21 December 2011)

Quantum fidelity between two density matrices $F(\rho_1,\rho_2)$ is usually defined as the trace of the operator $\mathcal{F} = \sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}$. We study the logarithmic spectrum of this operator, which we denote by the *fidelity spectrum*, in the cases of the XX spin chain in a magnetic field, a magnetic impurity inserted in a conventional superconductor, and a bulk superconductor at finite temperature. When the density matrices are equal, $\rho_1 = \rho_2$, the fidelity spectrum reduces to the entanglement spectrum. We find that the fidelity spectrum can be a useful tool in giving a detailed characterization of the different phases of many-body quantum systems.

DOI: 10.1103/PhysRevA.84.062318 PACS number(s): 03.67.−a, 05.70.Fh, 74.20.Fg

I. INTRODUCTION

A quantum system in a pure state is described by a density matrix which is just a projector onto that state. At zero temperature it is the projector to the ground state of the system. In general, the Hamiltonian of the system is a function of some parameters which determine the ground state. The quantum fidelity between two states (for two sets of parameters) is, in this simple case, the absolute value of the overlap between the ground states for the two sets of parameters. When the system is in a mixed state the density matrix is more complex. Typical situations that lead to mixed states are (i) reduced density matrices where a trace over some degrees of freedom is carried out, or (ii) systems at finite temperatures where the density matrix may be taken as the Boltzmann factor over the energy eigenstates.

In general the quantum fidelity [1] between two states characterized by two density matrices ρ_1 and ρ_2 may be defined as the trace of the fidelity operator $\mathcal{F}$

$$F(\rho_1,\rho_2) = \mathrm{Tr}\mathcal{F} = \mathrm{Tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}} \tag{1}$$

(see also the related work on the Bures metric [2], the transition probability [3], and the statistical distance between pure [4] and mixed [5] states). One may also consider the spectrum of the fidelity operator $\mathcal{F}(\rho_1,\rho_2)$. Its set of eigenvalues λ_i, which we denote the *fidelity operator spectrum*, and $-\ln\lambda_i$, which we call the *fidelity spectrum*, may provide more information as compared to the fidelity (its trace), in a way parallel to the extra information provided by the entanglement spectrum [6], as compared to the von Neumann entropy.

In the case of pure states $\rho_1 = |\mathrm{GS}_1\rangle\langle\mathrm{GS}_1|$ and $\rho_2 = |\mathrm{GS}_2\rangle\langle\mathrm{GS}_2|$ the fidelity is just the norm of the overlap

$$F(\rho_1,\rho_2) = |\langle\mathrm{GS}_1|\mathrm{GS}_2\rangle|. \tag{2}$$

If the two states are the same it is just the normalization of the state (taken as 1). In the case of two equal mixed states the fidelity is just

$$F(\rho,\rho) = \mathrm{Tr}\rho = 1 \tag{3}$$

and the operator $\mathcal{F}$ in this case has a set of eigenvalues $\lambda_i = r_i$ such that $-\ln r_i$ is called the entanglement spectrum, and has received considerable attention lately [6].

In this work we will analyze the fidelity spectrum for several physical systems paying particular attention to the vicinity of quantum phase transitions (QPT) as well as the properties characterizing their quantum phases.

The interplay between quantum information and condensed matter physics has been extensively considered using entanglement as a measure for the behavior of many-body systems [7]. The distinguishability between states has been used as a possible criterion to study quantum phase transitions [8]. By its own nature, the fidelity between the pure ground states signals a change of state as one approaches a quantum phase transition [9]. The fidelity between mixed states has also been used as a signature of quantum phase transitions [10,11] and to distinguish between different states of matter at finite temperatures [12–14] (for general considerations on fidelity and the Bures metric between mixed states, see Refs. [15–18]). A standard measure of entanglement in a system is the von Neumann entropy. However, as argued in Ref. [6], more information about a mixed state is obtained if the entanglement spectrum is analyzed. Considering reduced density matrices where part of the degrees of freedom are integrated over, such as dividing the system in real space into two parts A and B, it was shown in the context of the quantum Hall effect [6,19] and in the context of coupled spin chains [20] that the ground-state entanglement spectrum of A contains information about the excited energy states of the frontier of the subsystem A. In particular, in the quantum Hall effect the entanglement spectrum of the bulk system has a low-lying structure of levels that matches the edge states, and in the case of two coupled Heisenberg spin chains, considering the subsystem A as one of the chains, the entanglement spectrum has a structure that matches the energy excitations of a single Heisenberg chain. Other partitionings of the system have been proposed that lead to further information [21] and considering a partitioning in momentum space it was shown that information about energy excitations of a single Heisenberg chain is contained in the ground-state wave function through the entanglement spectrum [22].

An interesting connection between the fidelity operator spectrum and the entanglement spectrum can be established, via the so-called purification of mixed states and the Uhlmann theorem (see, for example, Ref. [23], page 410), which introduce the correspondence between density matrices and pure states in a larger Hilbert space.

 ©2011 American Physical Society

The purification of mixed states allows to interpret a mixed state $\rho = \sum_i r_i |r_i\rangle\langle r_i|$ as a partial state of a larger system in a global pure state $|\psi_\rho\rangle = \sum_i \sqrt{r_i}|r_i\rangle|r_i\rangle_A$. The system of interest S, given by the Hilbert space $\mathcal{H} = \text{span}\{|r_i\rangle\}$, is enlarged by a fictitious *ancilla* system A given by an equally dimensional Hilbert space $\mathcal{H}_A = \text{span}\{|r_i\rangle_A\}$, such that $\rho = \text{Tr}_A|\psi_\rho\rangle\langle\psi_\rho|$. Obviously, the basis $\{|r_i\rangle_A\}$ forming the ancilla's Hilbert space is not unique: any other basis $\{U_A|r_i\rangle_A\}$, with U_A being an arbitrary unitary operator, would suffice. Having two mixed states $\rho_1 = \sum_i r_i^{(1)}|r_i^{(1)}\rangle\langle r_i^{(1)}|$ and $\rho_2 = \sum_i r_i^{(2)}|r_i^{(2)}\rangle\langle r_i^{(2)}|$, two possible corresponding purifications are $|\psi_{\rho_1}\rangle = \sum_i \sqrt{r_i^{(1)}}|r_i^{(1)}\rangle|r_i^{(1)}\rangle_A$ and $|\psi_{\rho_2}\rangle = \sum_i \sqrt{r_i^{(2)}}|r_i^{(2)}\rangle|r_i^{(2)}\rangle_A$. Changing the bases of the ancilla system by choosing the arbitrary unitaries $U_A^{(1)}$ and $U_A^{(2)}$ changes the overlap $|\langle\psi_{\rho_1}|\psi_{\rho_2}\rangle|$, the change being determined by the unitary $U_A^{(12)} = (U_A^{(1)})^\dagger U_A^{(2)}$. The Uhlmann theorem states that the fidelity between two mixed states ρ_1 and ρ_2 is the maximum of the overlap between any purifications

$$F(\rho_1,\rho_2) = \max_{U_A^{(12)}} |\langle\psi_{\rho_1}|\psi_{\rho_2}\rangle| \equiv F(|\tilde{\psi}_{\rho_1}\rangle\langle\tilde{\psi}_{\rho_1}|, |\tilde{\psi}_{\rho_2}\rangle\langle\tilde{\psi}_{\rho_2}|), \tag{4}$$

where $|\tilde{\psi}_{\rho_1}\rangle$ and $|\tilde{\psi}_{\rho_2}\rangle$ represent particular pairs of purifications that satisfy the Uhlmann theorem achieving the equality between the two (mixed state and pure state) fidelities.

As a measure of global state distinguishability, fidelity can be considered as a general order parameter, encompassing all possible order parameters of the system considered. Thus, regions of nonanalyticity in the parameter space can signal out the system's critical behavior, without the need to consider particular order parameters and their associated correlation functions that display the triggering out of the phase transition. Using the above correspondence (4) between mixed states $\rho(q)$ and their purifications $|\tilde{\psi}_\rho(q)\rangle$ (note that the tilde refer to a particular choice of purifications $|\tilde{\psi}_\rho(q)\rangle$ that saturate Eq. (4) of the Uhlmann theorem), both parametrized by the same set of parameters q, it is possible to infer points of *quantum* phase transitions of the enlarged composite system consisting of our system of interest S and the ancilla A. Moreover, the entanglement spectrum of the enlarged $(S + A)$ system, with respect to this particular division into subsystems S and A, is given by the eigenvalues $\{r_i(q)\}$ of $\rho(q) = \sum_i r_i(q)|r_i(q)\rangle\langle r_i(q)|$.

In the case of mutually commuting mixed states $\rho(q)$ that have the common eigenbasis $|r_i(q)\rangle \equiv |r_i\rangle$, the (quantum) fidelity F is nothing but the classical fidelity F_c between two probability distributions $\{r_i(q)\}$ and $\{r_i(q')\}$, taken in two parameter points q and q'

$$F(\rho(q),\rho(q')) = \sum_i \lambda_i(q,q') = \sum_i \sqrt{r_i(q)r_i(q')} \equiv F_c(\{r_i(q)\},\{r_i(q')\}). \tag{5}$$

For this commuting case, the fidelity quantifies the distinguishability between two probability distributions (i.e., their change with respect to the change of parameters q). Analogously, the fidelity operator spectrum $\lambda_i(q,q')$ for mixed states $\rho(q)$ and $\rho(q')$ of the system S would quantify the change of the eigenvalues $r_i(q)$ of the reduced density matrix obtained tracing out the subsystem A from the enlarged system $(S + A)$, which is in the pure states $|\tilde{\psi}_\rho(q)\rangle$ and $|\tilde{\psi}_\rho(q')\rangle$ that saturate the Uhlmann theorem.

In a more general noncommuting case, the distinguishability between two mixed states will depend not only on the change of the eigenvalues of a system's mixed state $\rho(q)$, but also on the change of the corresponding eigenvectors. Therefore, the fidelity operator spectrum would carry the additional information regarding the rate of change of system's eigenvectors as well, which can be quantified by the Uhlmann geometric phase, the mixed-state generalization of the Berry phase, as shown in Ref. [12].

In the case of thermal equilibrium, the mixed state density matrix is given by the (normalized) Boltzmann factor $\rho = e^{-\beta H}/Z$, where $Z = \text{Tr}(e^{-\beta H})$ is the partition function, H the system's Hamiltonian, and $\beta = 1/(k_B T)$ (k_B being the Boltzmann constant and T the absolute temperature). This can also be written as $\rho = e^{-\beta(H-U+TS)} = e^{-[S-(U-H)/T]/k_B}$ since the partition function Z can be written in terms of the free energy $F = U - TS$ as $Z = e^{-\beta F}$, where $U = \langle H\rangle$ is the internal energy, and S the system's entropy. Note that $-F/T = S - U/T$ is nothing but the Légendre transformation of the entropy S with respect to the energy, called the Massieu function.

In the commuting case, the eigenbases of all density matrices are equal, and we are comparing the occupation probabilities of the states being considered in forming the mixture. In this case the fidelity operator for thermal states becomes $\mathcal{F} = \frac{e^{-\frac{1}{2}(\beta_1 H_1 + \beta_2 H_2)}}{\sqrt{Z_1 Z_2}}$ and its eigenvalues $\lambda_i = (Z_1 Z_2)^{-1/2} e^{-\frac{1}{2}(\beta_1 \varepsilon_i^{(1)} + \beta_2 \varepsilon_i^{(2)})} = e^{-\frac{1}{2k_B}(S_1 - \frac{U_1 - \varepsilon_i^{(1)}}{T_1} + S_2 - \frac{U_2 - \varepsilon_i^{(2)}}{T_2})}$ are given as simple functions of the energies $\varepsilon_i^{(1)}$ and $\varepsilon_i^{(2)}$ of the two mutually commuting Hamiltonians H_1 and H_2 and the corresponding Massieu functions.

The fidelity operator $\mathcal{F}$ can be studied using different basis states, associated with different representations, such as position, momentum, energy or charge, and spin. Rewriting the fidelity operator in these different representations allows us to look more directly at the specific relevant modes that are more directly associated with the usual order parameter, thus participating more actively in the critical phenomena accompanying the phase transition. In this way we can obtain a more complete and physical characterization of the phase transition and of its underlying physics mechanisms. While the entanglement spectrum has some relation to the energy spectrum of the edge states or even bulk states, the fidelity spectrum contains information about which eigenvalues have a larger contribution to the distinguishability between quantum states.

In this work we consider the fidelity spectrum of various systems. We start by considering two systems at zero temperature: a magnetic impurity inserted in a conventional superconductor and the XX chain in a magnetic field. In the first case, described in Sec. II, the magnetic impurity is coupled through a spin interaction to the spin density of the conduction electrons, tuned by a coupling J. As previously discussed [11,24,25], as the coupling J grows, the system goes through a first-order phase transition. At this point the

system becomes magnetized and various quantities such as local density of states, spin content, gap function, and quantum information measures can be used to detect this transition. For instance, various entanglement measures [26] and the partial state fidelity itself [11] have been used before. In the case of the XX chain describing spins $1/2$ confined to a plane (xy) and with a transverse magnetic field h, aligned along the z direction (Sec. III), there is a quantum phase transition from an XX phase, where the spins are aligned in the xy plane, if the magnetic field is small, and an Ising-like phase where the spins point along the field direction, if the field is strong enough. Considering the coupling between the spins as the energy scale, the second-order transition occurs at the point $h_c = 1$. This transition is also signaled in various ways such as the decrease of the fidelity near the critical point [8]. In Sec. IV we consider thermal states in a conventional superconductor.

II. MAGNETIC IMPURITY IN A SUPERCONDUCTOR

Consider first a classical spin immersed in a two-dimensional s-wave conventional superconductor. We use a lattice description of the system. In the center of the system $i = l_c = (0,0)$, we place a classical spin along the z direction $\vec{S} = S\vec{e}_z$, with no loss of generality. The Hamiltonian of the system is given by

$$\begin{aligned} H = &- \sum_{\langle ij\rangle\sigma} t_{ij} c^\dagger_{i\sigma} c_{j\sigma} - \varepsilon_F \sum_{i\sigma} c^\dagger_{i\sigma} c_{i\sigma} \\ &+ \sum_i (\Delta_i c^\dagger_{i\uparrow} c^\dagger_{i\downarrow} + \Delta^*_i c_{i\downarrow} c_{i\uparrow}) \\ &- \sum_{\sigma\sigma'} J c^\dagger_{l_c\sigma} \sigma^z_{\sigma\sigma'} c_{l_c\sigma'}, \end{aligned} \tag{6}$$

where the first term describes the hopping of electrons between different sites on the lattice, ε_F is the chemical potential, the third term is the superconducting s pairing with the site-dependent order parameter Δ_i, and the last term, with $J > 0$, is the exchange interaction between an electron at site $i = l_c$ and the magnetic impurity. The hopping matrix is given by $t_{ij} = t\delta_{j,i+\delta}$ where δ is a vector to a nearest-neighbor site. Note that both indices $i, j \in \{1,2,\ldots,N\}$ specify sites on a two-dimensional system (N is the number of sites). We take energy units in terms of t ($t = 1$), and choose $\varepsilon_F = -1$.

If we divide the whole system in two subsystems, say A and B, then the partial mixed state, given by the reduced density operator ρ_A for the subsystem A, is defined as

$$\rho_A = \mathrm{Tr}_B \rho, \tag{7}$$

where $\mathrm{Tr}_B[\cdot]$ represents the partial trace evaluated over the Hilbert space $\mathcal{H}_B$ of the subsystem B. We take A to be one site, either the impurity site or an arbitrary site in the bulk far from the impurity such that ρ_1 and ρ_2 of Eq. (1) are one-site density matrices.

In many-body systems second quantization is the natural way to perform any calculation. The matrix elements of the density matrix are simply defined in terms of the correlation functions of the whole system. For instance, in the case of the single-site partial mixed states using local basis states $\mathcal{B} = \{|0\rangle, |\uparrow\downarrow\rangle, |\uparrow\rangle, |\downarrow\rangle\}$, which denote the four possible states (unoccupied, double occupied, single occupied with an electron with spin up, and single occupied with an electron with spin down, respectively) it can be shown that the corresponding density matrix reads as [11]

$$\rho_i = \begin{pmatrix} \langle (1-n_\uparrow)(1-n_\downarrow)\rangle & \langle c^\dagger_\uparrow c^\dagger_\downarrow\rangle & 0 & 0 \\ \langle c_\downarrow c_\uparrow\rangle & \langle n_\uparrow n_\downarrow\rangle & 0 & 0 \\ 0 & 0 & \langle n_\uparrow(1-n_\downarrow)\rangle & \langle c^\dagger_\downarrow c_\uparrow\rangle \\ 0 & 0 & \langle c^\dagger_\uparrow c_\downarrow\rangle & \langle (1-n_\uparrow)n_\downarrow\rangle \end{pmatrix}_i, \tag{8}$$

where the index i denotes the site. The spin and the charge parts decouple. The spin part couples the two spin orientations (single occupied states) and the charge part couples the empty and doubly occupied states. The diagonal terms of the matrix describe the number of empty sites, the number of doubly occupied sites, the number of spin up sites, and the number of spin down sites, respectively. The sum of the diagonal terms is equal to 1 due to normalization. The matrix is easily diagonalized and the fidelity between two different one-site states obtained straightforwardly.

We will consider two density matrices of the form $\rho_1(J_1; i_1)$, and $\rho_2(J_2; i_2)$ for two values, J_1 and J_2, of the spin coupling J, and for two sites i_1 and i_2.

In Fig. 1 we show the fidelity operator spectrum as a function of the coupling between the magnetic impurity and the electronic spin density taking $J_1 = J$, $J_2 = J + 0.05$ where $i_1 = i_2 = i$ is the impurity site or a bulk site. The charge and spin parts separate and in the charge part there is the empty and the doubly occupied contributions and in the spin part the spin up and spin down contributions. As the coupling J grows there is a discontinuity in the eigenvalues. This discontinuity is associated with the quantum phase transition previously discussed. The sum of the four eigenvalues is the fidelity, as discussed before. There is a discontinuity at the QPT both in the total fidelity, in the charge and spin parts and in the individual eigenvalues as well. As we can see, the discontinuities occur mainly in one of the charge eigenvalues and in one of the spin eigenvalues. As one crosses the QPT to a regime where the impurity captures one electron breaking a Cooper pair we see that the main contribution to the discontinuity in the charge part comes from the doubly occupied states, where there is a significant increase at the QPT. In the same way there

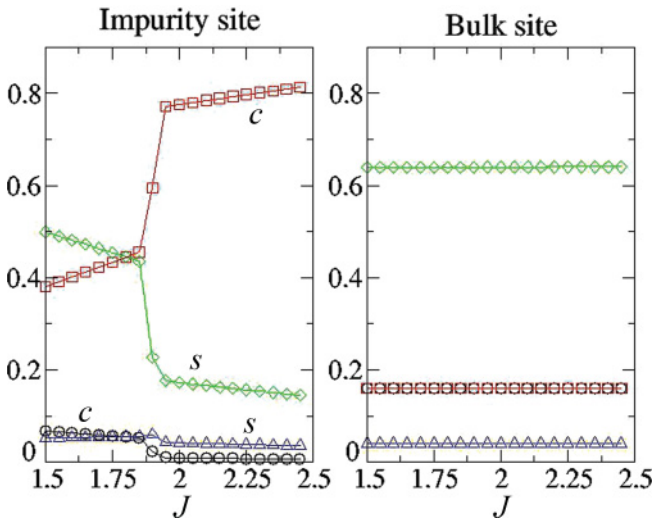


FIG. 1. (Color online) Fidelity operator spectrum at the impurity site (left) and at a bulk site (right) as a function of the spin coupling J, where one density matrix is calculated at J and the other at $J + \delta J$, where $\delta J = 0.05$. We show the charge contributions (c) and the spin contributions (s).

is a significant decrease in the spin up eigenvalue, leading to a smaller spin contribution beyond the QPT. For small values of J there is a screening of the perturbation induced by the magnetic impurity in the superconductor. A small fidelity means a higher degree of distinguishability. For small values of J the spin down eigenvalue is small but the spin up is still high. Beyond the QPT both contributions are small. This indicates that the transition is mainly of spin character: as the fidelity tends asymptotically to 1 away from the QPT and the spin eigenvalues are small, the charge eigenvalues have to compensate, however, mainly through the doubly occupied contribution, as expected. Note that, in the bulk, the transition is quite small. The physics is very local, centered around the impurity site. Finally, note the numerical differences between the relative drop of the total fidelity, which is for $\delta J = 0.05$ about 10% (see Fig. 1 from Ref. [11]), and the relative changes of the two most prominent modes, which is about 100%–150%.

In Fig. 2 we consider that the coupling is fixed, but the two density matrices are calculated at different sites i_1 and i_2: $\rho_1(J,i_1)$ and $\rho_2(J,i_2)$. The various cases are specified in the caption of the figure. In these figures the horizontal axis is the lattice site. For each site there are four eigenvalues. The central point is the impurity site. The symmetrical peaks close to the central point are the neighbors of the central point [please note that here we number the lattice sites of the two-dimensionl (2D) system as a one-dimensional (1D) system sequentially row by row; so the lattice nearest neighbors in the x direction are close by neighbors but the neighbors in the y direction are far apart). The fluctuations around the central peak are, of course, better seen in a three-dimensional (3D) plot. This is shown in Fig. 3 where we compare the four eigenvalues with the results of Fig. 1(a). As discussed above, the empty site and the spin down contributions are small and do not change much as we change from $J = 1.5$ (below the QPT) to $J = 2.5$ (above the QPT).

III. XX SPIN-1/2 CHAIN IN A TRANSVERSE FIELD

The XX spin 1/2 model has been solved exactly via the Jordan-Wigner transformation where it is reduced to a system of free spinless fermions [27]. The correlation functions were also calculated [28] as well as the reduced density matrix of a system of L contiguous spins [29–31]. The information theoretic approach in terms of fidelity, Fisher metric, and Chernoff bound was applied to the XY model in Refs. [8,13,14,32–37].

The Hamiltonian we will consider here is of the form

$$H = -\frac{1}{2}\sum_{l=0}^{N-1}\left(\frac{1+\gamma}{2}\sigma_l^x\sigma_{l+1}^x + \frac{1-\gamma}{2}\sigma_l^y\sigma_{l+1}^y + h\sigma_l^z\right). \tag{9}$$

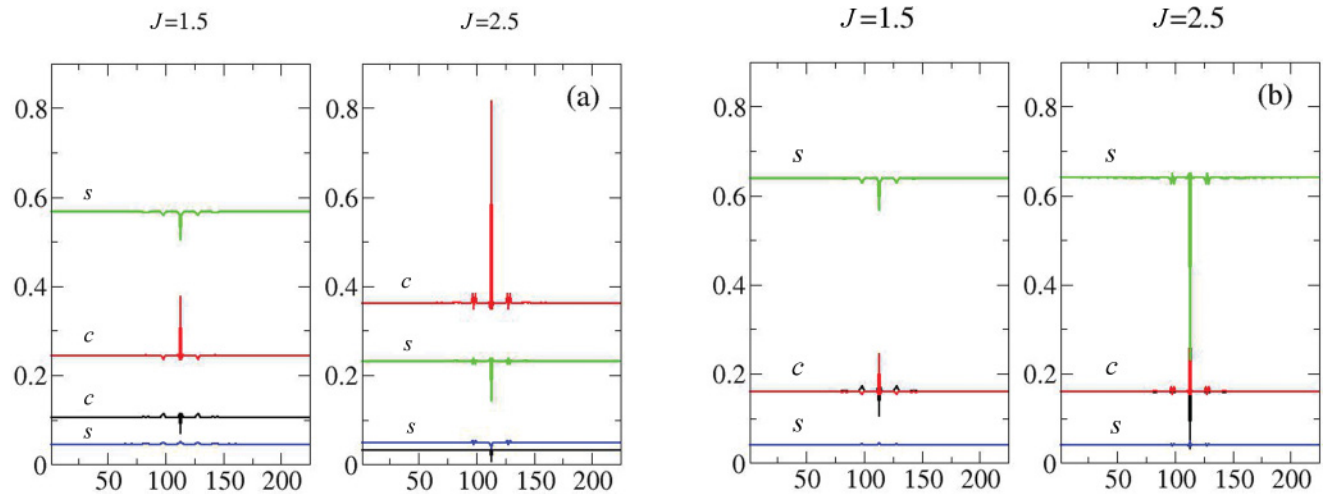


FIG. 2. (Color online) Fidelity operator spectrum: charge and spin eigenvalues. System size is 15×15. (a) One of the sites is at the impurity $i_1 = l_c$ and the other site i_2 is arbitrary. In the left panel $J_1 = J_2 = 1.5$, and in the right panel $J_1 = J_2 = 2.5$. (b) One of the sites, i_1, is a site in the bulk and the other site, i_2, is arbitrary. The other parameters are the same as in (a). Recall that $J = 1.5 < J_c$ and $J = 2.5 > J_c$.

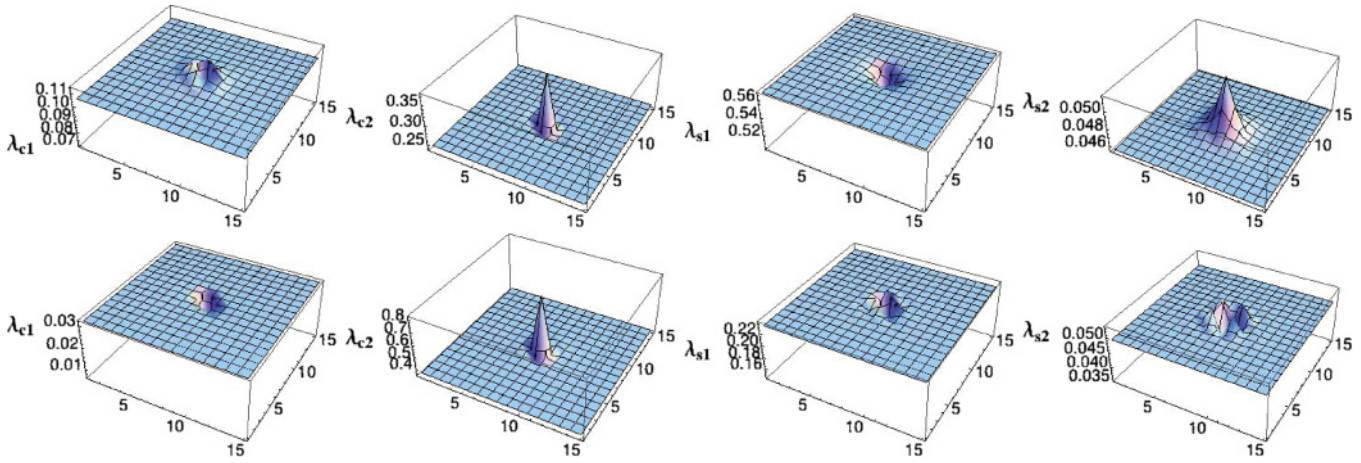


FIG. 3. (Color online) Fidelity operator spectrum as a function of real space in units of the lattice constant. System size is 15×15. Top panels $J = 1.5$. Lower panels $J = 2.5$. From left to right, the first two panels are the charge eigenvalues and the right two panels the spin eigenvalues, for lattice sites $i_1 = l_c$ and i_2 an arbitrary site.

Here the spin operators are described by Pauli matrices, h is the transverse magnetic field, and γ is the anisotropy. We will simplify and consider $\gamma = 0$ and we take $0 < h < 1$. We will consider a block of L contiguous spins. The reduced density matrix of the block can be written as [29]

$$\rho_A = \prod_{i=1}^{L} \left(\frac{1+\nu_i}{2} b_i^\dagger b_i + \frac{1-\nu_i}{2} b_i b_i^\dagger \right). \tag{10}$$

The operators b_i are spinless fermionic operators with a two-state space (the eigenvalues of the number operator $n_i = b_i^\dagger b_i$ are 1 and 0, corresponding to the occupied or empty state, respectively). The eigenvalues of the reduced density matrix are 2^L in number as the result of the direct product of the $i = 1, \ldots, L$ subspaces. Defining two Majorana operators of the form

$$c_{2l-1} = \left(\prod_{n=1}^{l-1} \sigma_n^z \right) \sigma_l^x, \quad c_{2l} = \left(\prod_{n=1}^{l-1} \sigma_n^z \right) \sigma_l^y, \tag{11}$$

in terms of the spin operators, it has been shown that

$$\langle \mathrm{GS}|c_m c_n|\mathrm{GS}\rangle = \delta_{m,n} + i(B_L)_{mn}. \tag{12}$$

The matrix B_L is written as

$$B_L = G_L \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{13}$$

with

$$G_L = \begin{pmatrix} g_0 & g_{-1} & \cdots & g_{1-L} \\ g_1 & g_0 & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ g_{L-1} & \cdot & \cdots & g_0 \end{pmatrix}. \tag{14}$$

We have that $g_l = g_{-l}$, $g_0 = (2\varphi_c/\pi) - 1$, and $g_{l\neq 0} = (2/l\pi)\sin l\varphi_c$, where $\varphi_c = \arccos(h)$. Defining the new Majorana fermions through the transformation $d = Vc$ and imposing that

$$\langle \mathrm{GS}|d_m d_n|\mathrm{GS}\rangle = \delta_{m,n} + i(\tilde{B}_L)_{mn} \tag{15}$$

with

$$\tilde{B}_L = V B_l V^T = \Omega \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{16}$$

where Ω is a diagonal matrix with L diagonal elements ν_l, leads to the diagonal form Eq. (10) of the reduced density matrix, ρ_A, having defined L complex fermionic fields like

$$b_l = \frac{d_{2l} + i d_{2l+1}}{2}. \tag{17}$$

The transformation V that block diagonalizes the problem depends on the Hamiltonian parameters through the numbers g_l.

To calculate the fidelity operator $\mathcal{F} = \sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}$ one needs to consider the product of two reduced density matrices for two different magnetic fields. Even though each of the reduced density matrices can be diagonalized, the transformations needed to diagonalize ρ_1 and ρ_2 are different. The diagonalization of ρ_1 (ρ_2) is obtained introducing a matrix V_1 (V_2). To obtain the spectrum of $\mathcal{F}$ we rewrite the expression for the diagonalized reduced density matrix ρ_2 in terms of the fermionic operators of the matrix ρ_1 which leads to $\rho_2 = e^{-H_2}$, where

$$H_2 = -\sum_{i=1}^{L}\sum_{j=1}^{L}\sum_{l=1}^{L} \left(\ln \frac{1+\nu_{2,l}}{2} T_{l,i} T_{l,j} |1\rangle_i \langle 0|_i \otimes |0\rangle_j \langle 1|_j + \ln \frac{1-\nu_{2,l}}{2} T_{l,i} T_{l,j} |0\rangle_i \langle 1|_i \otimes |1\rangle_j \langle 0|_j \right). \tag{18}$$

Here

$$T_{l,i} = \sum_{p=1}^{L} \bar{V}_{2,l,p} (\bar{V})^{-1}_{1,p,i}, \tag{19}$$

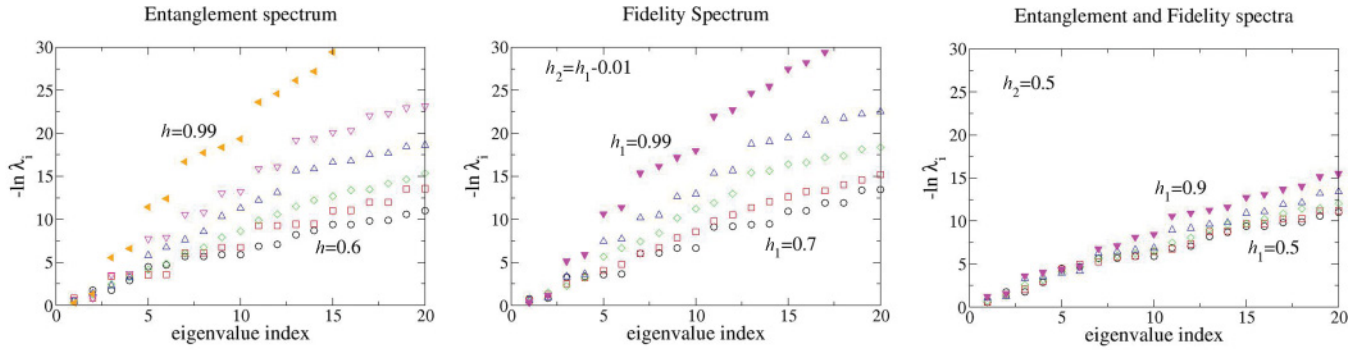


FIG. 4. (Color online) Fidelity spectrum $-\ln\lambda_i$, for system size $L = 6$. Left: Entanglement spectrum for different values of magnetic field. From bottom to top $h = 0.6,0.7,0.8,0.9,0.95,0.99$. Middle: Fidelity spectrum for $\delta h = 0.01$ and for $h = 0.7,0.8,0.9,0.95,0.99$. Right: Fidelity spectrum for pairs of the magnetic field as $h_1 = 0.5$, $h_2 = 0.5,0.6,0.7,0.8,0.9$. The magnetic field is in units of the spin coupling as in the following figures.

where we defined $\bar{V}$ by

$$V = \bar{V} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{20}$$

The diagonal elements $\nu_{2,l}$ are the diagonal elements of Ω for the magnetic field of ρ_2. Diagonalizing H_2, we obtain ρ_2, expressed in the eigenbasis of ρ_1, and obtain the spectrum of $\mathcal{F}$, as intended. Note that this requires diagonalizing a $2^L \times 2^L$ matrix which is much larger than the $L \times L$ matrix required for the entanglement entropy.

In Fig. 4 we show the fidelity spectrum $-\ln\lambda_i$. In the left panel we consider $\rho_1 = \rho_2$ and the operator $\mathcal{F}$ is just the density matrix. Therefore, its logarithmic spectrum is the entanglement spectrum. In the middle panel we consider the fidelity spectrum and in the right panel we compare the entanglement spectrum with the fidelity spectrum. As discussed before [22] there is no clear structure in the entanglement spectrum since we are considering a real space block of spins. The fidelity spectrum when we consider two very close values of the magnetic field is very similar to the entanglement spectrum. In both cases, as we increase the magnetic field and approach the critical point that separates the XX phase from the Ising phase, the entanglement or fidelity spectrum increases considerably. This implies that the fidelity operator spectrum is decreasing fast. The same can be seen when we compare the fidelity spectrum with the entanglement spectrum in the right panel. Here we are considering two values of the magnetic field h_1 and h_2 associated, respectively, with ρ_1 and ρ_2 that differ by a finite amount. Therefore there is a significant difference as the difference between h_1 and h_2 increases, even though we are far from the critical regime. The further the two points are the smaller the fidelity and the fidelity operator eigenvalues λ_i should be.

To analyze the spectrum we calculate the moments of the distribution of the eigenvalues [38] defined as

$$M_n = \sum_{i=1}^{2^L} \lambda_i^n. \tag{21}$$

The moment of order $n = 1$ is the fidelity. We also calculate the von Neumann entropy of the fidelity

$$S_1 = -\sum_{i=1}^{2^L} \lambda_i \ln\lambda_i \tag{22}$$

and the Rényi entropies

$$S_n = \frac{1}{1-n} \ln M_n. \tag{23}$$

Note that in the case of the entanglement ($\rho_1 = \rho_2$) the von Neumann entropy can also be obtained as

$$S_1 = -\sum_{i=1}^{L} \left(\frac{1+\nu_i}{2} \ln \frac{1+\nu_i}{2} + \frac{1-\nu_i}{2} \ln \frac{1-\nu_i}{2} \right) \tag{24}$$

which only involves a sum with L terms.

In Fig. 5 we present the first five moments for several choices of pairs of magnetic fields (h_1,h_2) and for several block sizes L. In the case of the entanglement spectrum, when $h_1 = h_2$, the first moment is just the trace of the density matrix (top left panel). When $h_1 \neq h_2$ it is the fidelity and therefore it is very close to 1 except in the vicinity of the phase transition where we consider $h_2 = h_1 - \delta h, \delta h = 0.01$ (top right panel). The higher momenta are very similar between the entanglement spectrum and the fidelity spectrum, in this case. A detailed study of the moments and Rényi entropies has been carried out in Ref. [39]. Considering two values of the magnetic fields further away from each other the moments change in structure. Fixing, for instance, $h_1 = 0.5$ (bottom left panel) or $h_1 = 0.98$ (bottom right panel) and varying h_2 we find that when h_2 crosses h_1 there is a sharp increase which equals 1 for the first moment (trivially since for this case the first moment is the trace of the density matrix). This maximum is also observed in the higher moments of the spectrum. Similar information can be obtained from the Rényi entropies. As explained above, S_1 is just the von Neumann entropy and the other entropies are proportional to the logarithms of the moments of the spectrum. In Fig. 6 we compare various cases for $L = 1,6$. As the magnetic field approaches the critical point we find that the various entropies have a minimum near the critical point showing that the partial state density matrix signals the QPT, as previously obtained. In the lower panels we consider

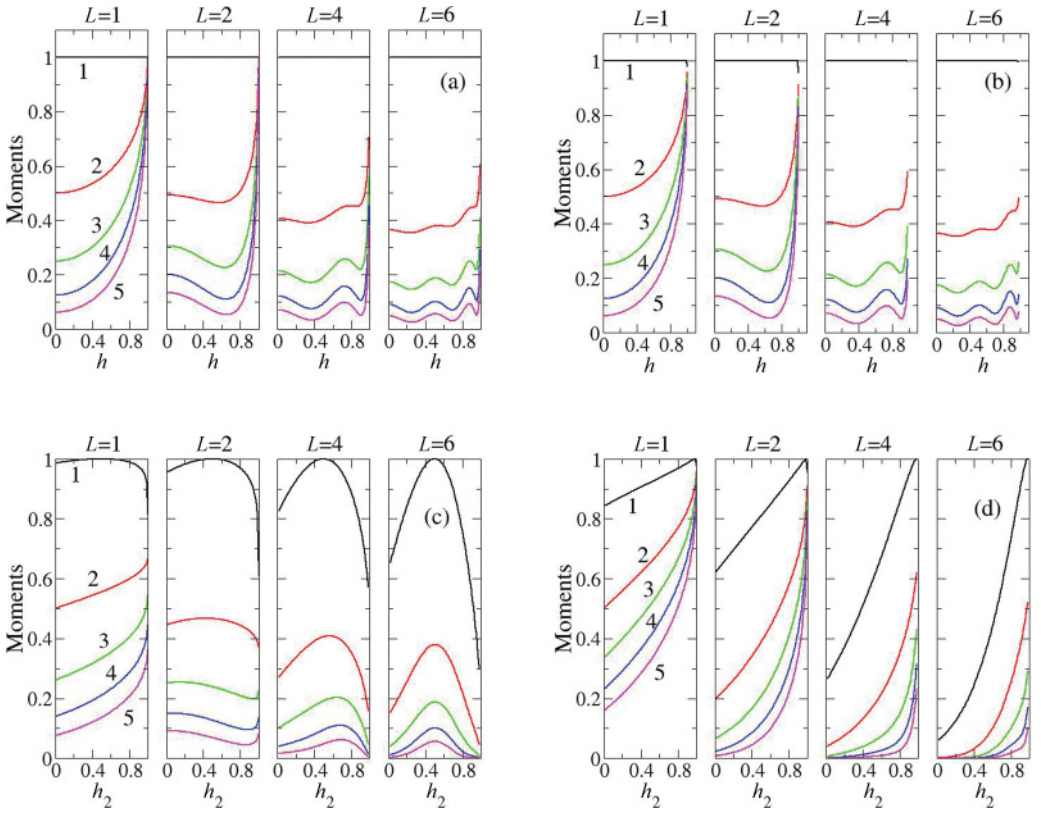


FIG. 5. (Color online) First five moments of the entanglement operator (density matrix) and fidelity operator spectra for blocks of sizes $L = 1,2,4,6$, as a function of magnetic field. In the top left panel we consider the moments of the entanglement operator spectrum. In the top right panel we consider the moments of the fidelity operator spectrum for two close-by values of h. In the bottom left (right) panel we consider the fidelity operator spectra for $h_1 = 0.5$ ($h_1 = 0.98$) and h_2 arbitrary.

the Rényi entropies associated with the fidelity operator. The structure is, in general, more complicated. Fixing $h_1 = 0.98$ close to the critical point, as h_2 approaches h_1 the entropies tend to those corresponding to the entanglement spectrum. Far from this point the entropies differ considerably and depend strongly on the block size. For instance for $L = 1$ they are of the order of the entropies for the entanglement spectrum, but for $L = 6$ the entropies are considerably higher, except the von Neumann entropy. Also, note that the Rényi entropy S_1 in the fidelity case has a depression at small magnetic fields.

We also calculate the fidelity susceptibility, introduced in Refs. [34,40], with its geometrical meaning discussed in Refs. [34,41]. Global fidelity susceptibility for the XY model was discussed in Refs. [8,34]. In the present paper we calculate the fidelity susceptibility associated with the reduced density matrices of blocks of spins. It is defined as

$$\chi_F = \sum_{i=1}^{2^L} \chi_{F,i}, \quad \chi_{F,i} = \frac{\partial^2 \lambda_i}{\partial(\delta h)^2}. \tag{25}$$

In Fig. 7 we present the fidelity susceptibility associated with blocks of different sizes as a function of the magnetic field. The divergence of the susceptibility is clearly seen as we approach the QPT.

IV. THERMAL STATES OF A CONVENTIONAL BCS SUPERCONDUCTOR

In this section we consider a conventional s-wave superconductor at finite temperature described by the effective mean-field BCS Hamiltonian

$$H_{\rm BCS}^{\rm eff} = \sum_k \varepsilon_k (n_{k\uparrow} + n_{-k\downarrow}) - \sum_k (\Delta_k c_{k\uparrow}^\dagger c_{-k\downarrow}^\dagger + \Delta_k^* c_{-k\downarrow} c_{k\uparrow} - \Delta_k^* \langle c_{-k\downarrow} c_{k\uparrow} \rangle), \tag{26}$$

with $\Delta_k = -V\langle c_{-k\downarrow} c_{k\uparrow}\rangle$, where the lattice-mediated pairing interaction is constant and nonvanishing between electrons around the Fermi level only. The density matrix is given by [12]

$$\rho = \frac{1}{Z} e^{-(H_{\rm BCS}^{\rm eff} - \mu N)/T} = \frac{e^{\sum_k \vec{\tilde{h}}_k \vec{T}_k + K}}{{\rm Tr}[e^{\sum_k \vec{\tilde{h}}_k \vec{T}_k + K}]} = \frac{\prod_k e^{\vec{\tilde{h}}_k \vec{T}_k}}{\prod_k {\rm Tr}[e^{\vec{\tilde{h}}_k \vec{T}_k}]}, \tag{27}$$

where T is the temperature, $\vec{\tilde{h}}_k = (\tilde{h}_k^+, \tilde{h}_k^-, \tilde{h}_k^0) = (2\Delta_k^*/T, 2\Delta_k/T, -2\bar{\varepsilon}_k/T)$, $\vec{T}_k = (T_k^+, T_k^-, T_k^0)$, $K = -1/T \sum_k (\bar{\varepsilon}_k + \Delta_k^* b_k)$ and $\bar{\varepsilon}_k = \varepsilon_k - \mu$. The norms of the vectors $\vec{\tilde{h}}_k$ are given by $\tilde{h}_k = 2E_k/T$, with $E_k = \sqrt{\bar{\varepsilon}_k^2 + |\Delta_k|^2}$. The coefficients $\vec{\tilde{h}}_k = \vec{\tilde{h}}_k(T,V)$ are functions of both the coupling constant V and the temperature T, through the gap parameters $\Delta_k = \Delta_k(T,V)$ and the chemical potential μ. By $n_{k\sigma} = c_{k\sigma}^\dagger c_{k\sigma}$ we denote the one-particle number operators,

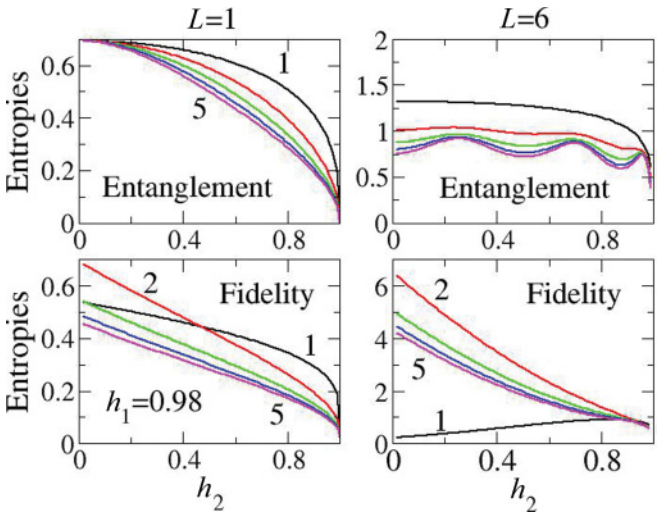


FIG. 6. (Color online) Rényi and von Neumann entropies for $L = 1$ (left) and $L = 6$ (right), for the first five moments, as a function of magnetic field. In the top panels we consider the entanglement operator spectrum and in the bottom panels the fidelity operator spectrum as a function of h_2, for $h_1 = 0.98$. Note the decrease of the entropies close to the quantum critical point, both for the entanglement case and the fidelity case.

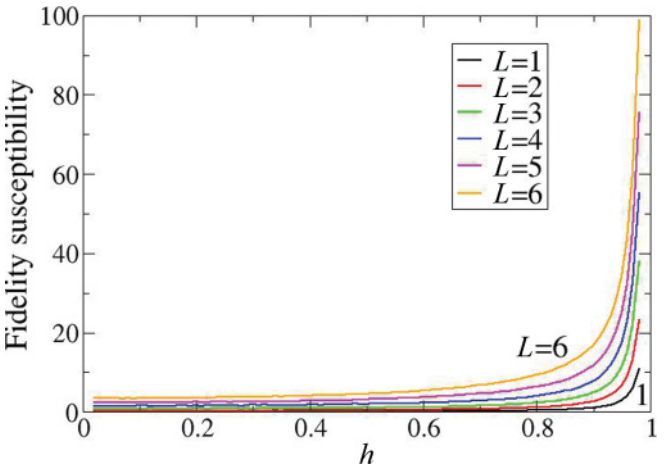


FIG. 7. (Color online) Block fidelity susceptibility as a function of magnetic field for $L = 1, \ldots, 6$ which shows the sharp increase near the critical point.

while by $b_k^\dagger = c_{k\uparrow}^\dagger c_{-k\downarrow}^\dagger$ and $b_k = c_{-k\downarrow} c_{k\uparrow}$. The T_k operators are given by $T_k^+ = b_k^\dagger$, $T_k^- = b_k$ and $2T_k^0 + 1 = (n_{k\uparrow} + n_{-k\downarrow})$ and form an su(2) algebra.

The fidelity is given by

$$\begin{aligned} F(\rho_a, \rho_b) &= \mathrm{Tr}\left[\left(\rho_a^{1/2}\rho_b\rho_a^{1/2}\right)^{1/2}\right] \\ &= \frac{\mathrm{Tr}\left[\left(\prod_k e^{\frac{\vec{a}_k}{2}\vec{T}_k} e^{\vec{b}_k\vec{T}_k} e^{\frac{\vec{a}_k}{2}\vec{T}_k}\right)^{1/2}\right]}{\prod_k (\mathrm{Tr}[e^{\vec{a}_k\vec{T}_k}]\mathrm{Tr}[e^{\vec{b}_k\vec{T}_k}])^{1/2}} \\ &= \mathrm{Tr}\left[\prod_k (\mathcal{F}_k)^{1/2}\right]. \end{aligned} \tag{28}$$

As for every k the operators $\vec{T}_k$ form an su(2) algebra, and therefore by exponentiation define a Lie group, we can write $e^{\frac{\vec{a}_k}{2}\vec{T}_k} e^{\vec{b}_k\vec{T}_k} e^{\frac{\vec{a}_k}{2}\vec{T}_k} = e^{2\vec{c}_k\vec{T}_k}$. For each value of the momentum we get a four-dimensional space with respect to momentum states that are empty, doubly occupied, or singly occupied by a spin up or a spin down electron. The space is therefore of the type $\mathcal{B} = \{|0\rangle, |\uparrow\downarrow\rangle, |\uparrow\rangle, |\downarrow\rangle\}$, similarly to the problem of the impurity in a superconductor. The fidelity operator is then easily diagonalized in this 4×4 subspace. We will study the possible eigenvalues of the fidelity operator for each momentum. As before, the charge and spin parts separate. Moreover, in this problem the two spin components are degenerate. Therefore it is enough to look at three eigenvalues (two for the charge part and one for the spin part). The results are presented in Figs. 8 through 12 and discussed in more detail below.

The fidelity operator for each momentum value (denoted the *k-fidelity operator*) is of the form $\mathcal{F}_k = \sqrt{A_k}$, where

$$A_k = \frac{1}{D_k}\begin{pmatrix} \alpha_k & \beta_k & 0 & 0 \\ \beta_k & \gamma_k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{29}$$

The matrix elements are given by

$$\begin{aligned} \alpha_k &= \cosh E_k^a/T \cosh E_k^b/T + \sinh E_k^a/T \sinh E_k^b/T \frac{\Delta_k^a\Delta_k^b + \bar{\epsilon}_k^a\bar{\epsilon}_k^b}{E_k^a E_k^b} + \sinh E_k^a/T \cosh E_k^b/T \frac{\bar{\epsilon}_k^a}{E_k^a} + \sinh E_k^b/T \frac{\bar{\epsilon}_k^b}{E_k^b} \\ &\quad + \left(\cosh E_k^a/T - 1\right) \sinh E_k^b/T \frac{\Delta_k^a\Delta_k^b + \bar{\epsilon}_k^a\bar{\epsilon}_k^b}{E_k^a E_k^b} \frac{\bar{\epsilon}_k^a}{E_k^a}, \end{aligned} \tag{30}$$

$$\begin{aligned} \gamma_k &= \cosh E_k^a/T \cosh E_k^b/T + \sinh E_k^a/T \sinh E_k^b/T \frac{\Delta_k^a\Delta_k^b + \bar{\epsilon}_k^a\bar{\epsilon}_k^b}{E_k^a E_k^b} - \sinh E_k^a/T \cosh E_k^b/T \frac{\bar{\epsilon}_k^a}{E_k^a} - \sinh E_k^b/T \frac{\bar{\epsilon}_k^b}{E_k^b} \\ &\quad - \left(\cosh E_k^a/T - 1\right) \sinh E_k^b/T \frac{\Delta_k^a\Delta_k^b + \bar{\epsilon}_k^a\bar{\epsilon}_k^b}{E_k^a E_k^b} \frac{\bar{\epsilon}_k^a}{E_k^a} \end{aligned} \tag{31}$$

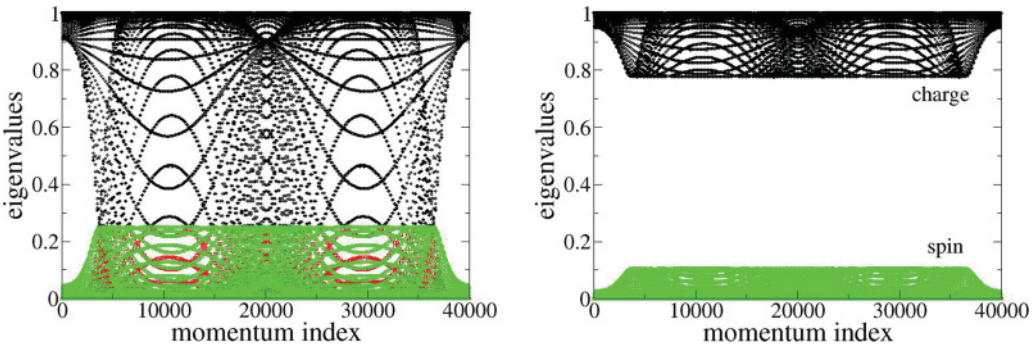


FIG. 8. (Color online) Exponential entanglement spectrum for a bulk superconductor as a function of momentum, labeled sequentially row by row in the Brillouin zone, in the normal phase ($\Delta = 0$, left panel), and in the superconducting phase ($\Delta \neq 0$, right panel), at temperature T. Note that in the right panel the lowest charge eigenvalue is smaller than the spin eigenvalues.

$$\beta_k = \sinh E_k^a/T \cosh E_k^b/T \frac{\Delta_k^a}{E_k^a} + \sinh E_k^b/T \frac{\Delta_k^b}{E_k^b} + \left(\cosh E_k^a/T - 1\right) \sinh E_k^b/T \frac{\Delta_k^a \Delta_k^b + \bar{\epsilon}_k^a \bar{\epsilon}_k^b}{E_k^a E_k^b} \frac{\Delta_k^a}{E_k^a}, \tag{32}$$

and

$$D_k = 2\left(1 + \cosh E_k^a/T\right)2\left(1 + \cosh E_k^b/T\right). \tag{33}$$

The eigenvalues of the k-fidelity operator $\mathcal{F}_k$ are therefore of the form

$$\frac{1}{\sqrt{D_k}} \begin{pmatrix} \eta_+^k & 0 & 0 & 0 \\ 0 & \eta_-^k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{34}$$

where

$$\eta_\pm^k = \frac{1}{2}\left[(\alpha_k + \gamma_k) \pm \sqrt{(\alpha_k - \gamma_k)^2 + 4\beta_k^2}\right]. \tag{35}$$

We will be interested in situations where ρ_1 and ρ_2 correspond to points in parameter space, which we choose to be the temperature T and the gap function Δ_k that are far apart and may be in the same or different thermodynamic phases.

In Fig. 8 we present the k-fidelity operator spectrum (of the operator $\mathcal{F}_k$) for the case when the two density matrices are equal ("exponential entanglement spectrum") where we compare the system in the normal phase (left) with the superconducting phase (right). The horizontal axis is an index over the eigenvalues and in the vertical axis we plot λ_k. We only plot three eigenvalues because the spin eigenvalues are degenerate. For each label the sum over the four eigenvalues is 1 due to normalization. In both phases the higher eigenvalue is the charge eigenvalue corresponding to empty sites (this will be discussed later on). In the normal phase the lowest eigenvalues merge into the higher eigenvalues, but in the superconducting phase the energy gap is clearly visible. Note that the eigenvalues are now labeled by the momentum. There is no partitioning of the system in real space, but there is a partitioning of the system in momentum space (since the system can be block diagonalized). Recall, however, that here the mixed state originates in the thermal states. We stress that we are not plotting the fidelity operator $\mathcal{F} = \prod_k \mathcal{F}_k$ eigenvalues. In Fig. 9 we consider two different density matrices where we plot the k-fidelity operator spectrum where one of the density matrices corresponds to a point in phase space in the normal phase and the other in the superconducting phase. In the left panel the temperatures T_a, T_b are the same (this can be obtained, for instance, considering two coupling constants) and in the right panel the temperatures T_a, T_b are different. In the first case the gap is still clearly visible. When the temperatures are different the gap remains the same. There

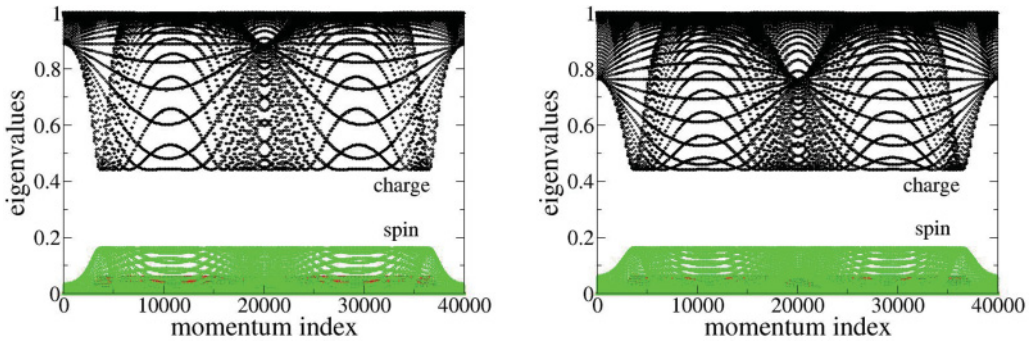


FIG. 9. (Color online) Fidelity operator spectrum for a bulk superconductor as a function of momenta, between the normal phase and the superconducting phase, at the same temperature (left) and at different temperatures (right).

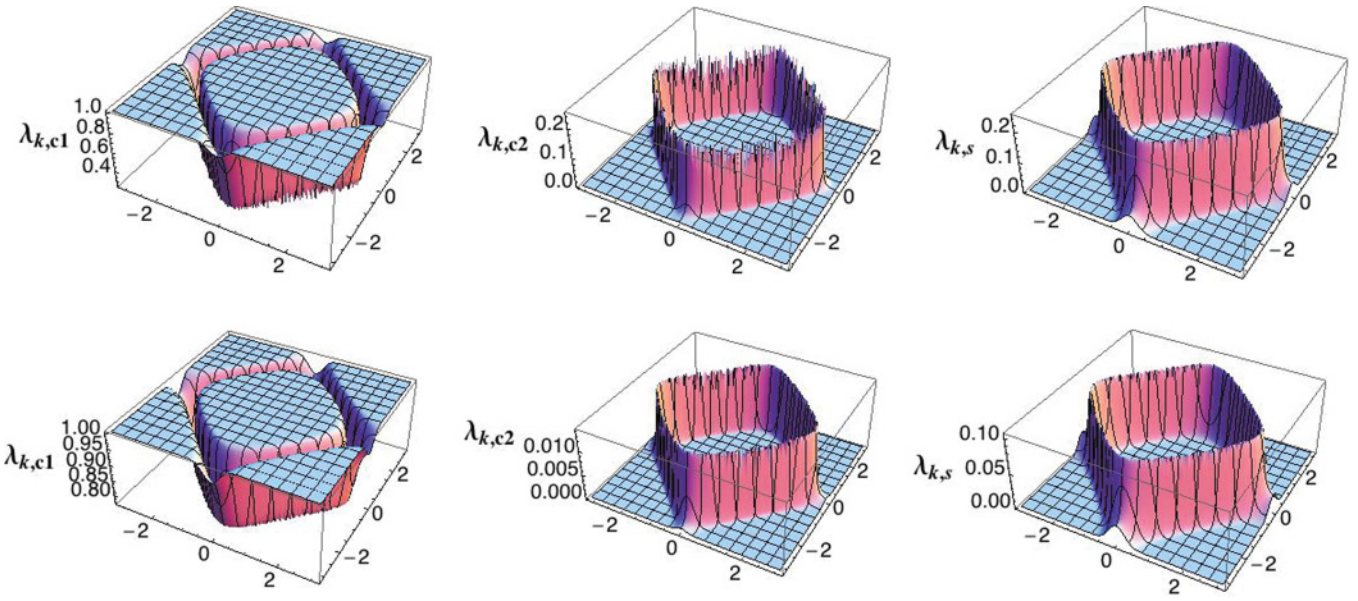


FIG. 10. (Color online) Top (bottom) panels: Normal (superconducting) phase exponential entanglement spectrum as a function of momentum k_x,k_y. In the three panels we show three eigenvalues since the two spin eigenvalues are degenerate.

is a small decrease of the highest charge eigenvalue that can be traced to the vicinity of the Fermi surface, as shown ahead.

To understand the spectrum in greater detail we consider the various eigenvalues in momentum space. In Fig. 10 we consider the system in the normal phase and in the superconducting phase for two equal density matrices $\rho_a = \rho_b$ and plot the three eigenvalues of the k-fidelity operator as a function of momentum for a 2D system. The fidelity is one for all momenta since $\mathrm{Tr}\rho_a = \mathrm{Tr}\rho_b = 1$. The depression in the highest eigenvalue (left panels) marks clearly the Fermi surface. For momenta larger than the Fermi momentum there are no electrons (except for thermal excitations contained in the Fermi function). So the eigenvalue corresponding to empty states is 1. The other eigenvalues are close to zero outside the Fermi surface. Due to the particle-hole transformation of the Bogoliubov transformation, the empty site eigenvalue (corresponding to doubly occupied sites in terms of the electrons) is also close to 1 inside the Fermi surface. Accordingly, the other eigenvalues are also close to zero inside the Fermi surface. The noticeable features are therefore close to the Fermi surface. A similar structure is observed in the superconducting phase. There is a slight change close to the Fermi surface which is due to the opening of the superconducting gap. The amplitude of the doubly occupied and spin eigenvalues are smaller in this case since the decrease of the first eigenvalue is smaller along the Fermi surface.

In Fig. 11 we show the fidelity operator spectrum for the case when one density matrix is in the normal phase and the other corresponds to a quantum state in the superconducting phase. The last panel shows the total fidelity. It is significantly decreased around the Fermi surface where the difference between the normal phase and the superconducting phase is larger due to the pairing and opening of the gap. Note that the highest eigenvalue has a structure that strongly resembles the total fidelity, but the numerical decrease around the Fermi surface is significantly more prominent than in the case of total fidelity (about 0.4, with respect to about 0.15 for total fidelity).

Finally, in Fig. 12 we compare the total fidelity as a function of momentum for different temperatures. The right panel corresponds to two quantum states in the normal phase, but at different temperatures. As expected, around the Fermi surface the fidelity decreases, however, there is a sharp region where it approaches 1. The width of the region around the Fermi surface is determined by the temperature through the Fermi function. The sharp maximum corresponds to the point where the two Fermi functions cross and so it pinpoints the location of the Fermi surface.

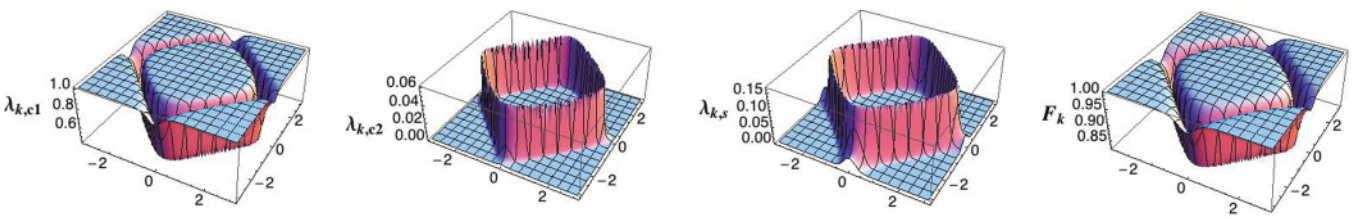


FIG. 11. (Color online) Fidelity operator spectrum between superconducting and normal phases, as a function of momentum. In the first three panels we show the three eigenvalues and in the last one the fidelity (smaller than 1 close to the Fermi surface). Note that the temperature is the same, with Δ either finite (superconducting phase) or zero (normal phase).

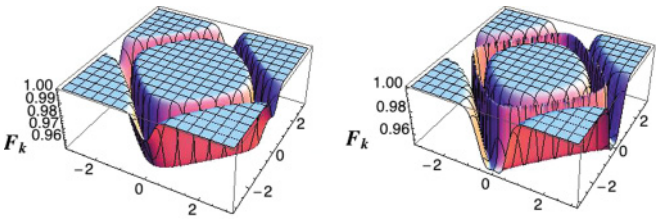


FIG. 12. (Color online) Total fidelity as a function of momentum. Left panel: Different temperatures and same finite Δ; right panel: different temperatures and $\Delta = 0$.

V. CONCLUSION

We have introduced and analyzed the fidelity spectrum and the fidelity operator spectrum for partial states for different systems such as a magnetic impurity in a conventional superconductor, a XX spin-$1/2$ chain in a transverse magnetic field, and the thermal states of a finite temperature bulk superconductor.

In the first problem we have found that only one charge eigenvalue and one spin eigenvalue have important changes as the quantum phase transition induced by the magnetic impurity occurs. The transition is associated with the capture of one electron by the impurity with a parallel spin. This feature is clearly seen when we consider two density matrices associated with the same lattice site (particularly the impurity site) and different but close by spin couplings between the impurity and the spin density of the electrons. Selecting two density matrices with the same spin coupling but different lattice sites leads to a signature of the phase transition that can be seen both from the eigenvalues associated with lattice sites far from each other and from the same lattice site.

In the spin chain problem we have studied the block fidelity and the block fidelity susceptibility and found that the quantum phase transition that occurs between an XX phase and an Ising-like phase is well signaled by the block fidelity. Both the entanglement spectrum and the fidelity spectrum do not show any significant features and we analyzed the spectra calculating the moments of the distribution and the Rényi entropies. The S_1 Rényi entropy associated with the block fidelity shows a distinctive characteristic away from the critical point.

Finally, in the finite temperature bulk superconductor we showed that in the superconductor there is a clear gap between the various k eigenvalues, as for the energy spectrum. In the case of two different density matrices we found that the effect of temperature is stronger than the difference in the order parameter distinguishing the normal from the superconducting phase. Analysing the k-fidelity operator spectrum it was clearly seen that the properties are determined by the structure around the Fermi energy, as expected. In the case of two density matrices for different temperatures in the normal phase the fidelity has a sharp maximum at the location of the Fermi surface, determined by the crossing of the Fermi functions.

We have shown that the fidelity spectrum, which we have introduced, can give a more detailed description and characterization of the phase transitions of many-body quantum systems providing complementary information to other techniques. Moreover, it allows to identify the modes that have a larger contribution to the distinguishability between the states across a phase transition and singles out the most divergent contributions to the susceptibility. Therefore, we hope that this can be applied to nontrivial problems where the traditional Ginzburg-Landau theory with a local order parameter is not known.

Note added: Recently another work appeared where the importance of fidelity modes is stressed [42].

ACKNOWLEDGMENTS

NP thanks the project of SQIG at IT, funded by FCT and EU FEDER Projects No. QSec PTDC/EIA/67661/2006, No. QuantPrivTel PTDC/EEA-TEL/103402/2008, No. FCT PEst-OE/EEI/LA0008/2011, and IT Project QuantTel, as well as the Network of Excellence, Euro-NF. VRV and PDS thank the Project No. PTDC/FIS/64926/2006, of Fundação para a Ciência e a Tecnologia, Portugal.

[1] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).
[2] D. Bures, Trans. Amer. Math. Soc. **135**, 199 (1969).
[3] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
[4] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
[5] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[6] H. Li and F. D. M. Haldane, Phys. Rev. Lett. **101**, 010504 (2008).
[7] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. **80**, 517 (2008).
[8] P. Zanardi and N. Paunković, Phys. Rev. E **74**, 031123 (2006).
[9] S. J. Gu, Int. J. Mod. Phys. B **24**, 4371 (2010).
[10] H. Q. Zhou, e-print arxiv:0704.2945.
[11] N. Paunković, P. D. Sacramento, P. Nogueira, V. R. Vieira, and V. K. Dugaev, Phys. Rev. A **77**, 052302 (2008).
[12] N. Paunković and V. R. Vieira, Phys. Rev. E **77**, 011129 (2008).
[13] P. Zanardi, H. T. Quan, X. Wang, and C. P. Sun, Phys. Rev. A **75**, 032109 (2007).
[14] P. Zanardi, L. Campos Venuti, and P. Giorda, Phys. Rev. A **76**, 062318 (2007).
[15] J. Twamley, J. Phys. A **29**, 3723 (1996).
[16] H. Scutaru, J. Phys. A **31**, 3659 (1998).
[17] G.-S. Paraoanu and H. Scutaru, Phys. Rev. A **58**, 869 (1998).
[18] H.-J. Sommers and K. Życzkowski, J. Phys. A **36**, 10083 (2003).
[19] N. Regnault, B. A. Bernevig, and F. D. M. Haldane, Phys. Rev. Lett. **103**, 016801 (2009).
[20] D. Poilblanc, Phys. Rev. Lett. **105**, 077202 (2010).
[21] A. Sterdyniak, N. Regnault, and B. A. Bernevig, Phys. Rev. Lett. **106**, 100405 (2011).
[22] R. Thomale, D. P. Arovas, and B. A. Bernevig, Phys. Rev. Lett. **105**, 116805 (2010).
[23] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[24] A. Sakurai, Prog. Theor. Phys. **44**, 1472 (1970).

[25] A. V. Balatsky, I. Vekhter, and J.-X. Zhu, Rev. Mod. Phys. **78**, 373 (2006).
[26] P. D. Sacramento, P. Nogueira, V. R. Vieira, and V. K. Dugaev, Phys. Rev. B **76**, 184517 (2007).
[27] E. Lieb, T. Schultz, and D. Mattis, Ann. Phys. (NY) **16**, 407 (1961).
[28] E. Barouch and B. McCoy, Phys. Rev. A **3**, 786 (1971).
[29] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003).
[30] J. I. Latorre, E. Rico, and G. Vidal, Quant. Inf. Comput.**4**, 48 (2004).
[31] B.-Q. Jin and V. E. Korepin, J. Stat. Phys. **116**, 79 (2004).
[32] P. Zanardi, M. Cozzini, and P. Giorda, J. Stat. Mech. **2**, L02002 (2007).
[33] M. Cozzini, P. Giorda, and P. Zanardi, Phys. Rev. B **75**, 014439 (2007).
[34] P. Zanardi, P. Giorda, and M. Cozzini, Phys. Rev. Lett. **99**, 100603 (2007).
[35] D. F. Abasto, N. T. Jacobson, and P. Zanardi, Phys. Rev. A **77**, 022327 (2008).
[36] S. Garnerone, N. T. Jacobson, S. Haas, and P. Zanardi, Phys. Rev. Lett. **102**, 057205 (2009).
[37] J. Dubail and J.-M. Stéphan, J. Stat. Mech. **3**, L03002 (2011).
[38] F. Pollmann and J. E. Moore, New J. Phys. **12**, 025006 (2010).
[39] A. R. Its and V. E. Korepin, J. Stat. Phys. **137**, 1014 (2009).
[40] W. L. You, Y. W. Li, and S. J. Gu, Phys. Rev. E **76**, 022101 (2007).
[41] L. Campos Venuti and P. Zanardi, Phys. Rev. Lett. **99**, 095701 (2007).
[42] S.-J. Gu, W.-C. Yu, and H.-Q. Lin, e-print arXiv:1108.2832.

Fidelity between partial states as a signature of quantum phase transitions

N. Paunković,[1] P. D. Sacramento,[2] P. Nogueira,[2] V. R. Vieira,[2] and V. K. Dugaev[2,3]
[1]*SQIG–Instituto de Telecomunicações, IST, Technical University of Lisbon (TULisbon), P-1049-001 Lisbon, Portugal*
[2]*CFIF and Department of Physics, IST, Technical University of Lisbon (TULisbon), P-1049-001 Lisbon, Portugal*
[3]*Department of Physics, Rzeszów University of Technology, Al. Powstańców Warszawy 6, 35-959 Rzeszów, Poland*
(Received 6 February 2008; published 1 May 2008)

We introduce a partial state fidelity approach to quantum phase transitions. We consider a superconducting lattice with a magnetic impurity inserted at its center, and look at the fidelity between partial (either one-site or two-site) quantum states. In the vicinity of the point of the quantum phase transition, we observe a sudden drop of the fidelity between two one-site partial states corresponding either to the impurity location or its close vicinity. This enables us to identify the on-site magnetization as the order parameter for the phase transition studied. In the case of two-site states, the fidelity reveals the transition point as long as one of the two electron sites is located at the impurity, while the other lies elsewhere in the lattice. We also determine the Uhlmann mixed state geometric phase, recently introduced in the study of the structural change of the system state eigenvectors in the vicinity of the lines of thermal phase transitions, and find it to be trivial, both for one- and two-site partial states, except when an electron site is at the impurity. This means that the system partial state eigenvectors do not contribute significantly to the enhanced state distinguishability around the point of this quantum phase transition. Finally, we use the fidelity to analyze the total amount of correlations contained within a composite system, showing that, even for the smallest one-site states, it features an abrupt quantitative change in the vicinity of the point of the quantum phase transition.

DOI: 10.1103/PhysRevA.77.052302 PACS number(s): 03.67.−a, 05.70.Fh, 75.40.Cx

I. INTRODUCTION

Recently, considerable developments have been made within the research born out of the collaboration between physicists from two rather distinct fields, those of quantum information and computation and those of condensed matter and many-body physics. Most of this interdisciplinary work has focused on the realization of quantum-information protocols with real many-body systems (for an overview, see [1]). Yet, a considerable development has been accomplished by moving in the "opposite" direction, i.e., applying concepts and techniques widely used within the field of quantum information and computation to the study of problems relevant within the field of condensed matter physics. One of such topics, recently researched to a great extent, is that of the use of entanglement in the study of zero-temperature quantum phase transitions [2], as well as general thermal phase transitions (for an overview, see [3], Chaps. 19–21, and references therein).

One application of the quantum-information techniques to problems dealing with macroscopic many-body systems is the use of the (ground-) state fidelity [4] in studying (quantum-) phase transitions. It was first reported in [5], where the sudden drop of the fidelity along the regions of quantum phase transitions was observed in the Dicke and the *XY* models. Consequently, this method was applied to the cases of free Fermi systems [6], Bose-Hubbard models [7,8], spin chains [9,10], and in connecting the fidelity with the renormalization group flows [11]. Also, the cases of nonsymmetry breaking quantum phase transitions that are not of the Landau-Ginzburg-Wilson type were studied [12–14]. A more formal and general study based on differential geometric insights, establishing a connection between quantum phase transitions and the relevant thermodynamic response functions, was developed in [15,16] (see also [17]). Finally, using the mixed state fidelity, and extending the parameter space to include the temperature, quantum phase transitions were studied in [18,19], while the thermal phase transitions were explored in [15,17,20]. It is also worth noting here that the recently studied connection between the quantum phase transitions and Berry geometric phases [21] was established in [15,16] on a more formal level, unifying the fidelity and the Berry phase approaches to quantum phase transitions. Subsequently, the Uhlmann mixed state geometric phase [22] (the generalization of the pure state Berry phase) was introduced in [20] as a tool in exploring the change of the system thermal equilibrium state eigenvectors in the vicinity of the line of a phase transition.

In this paper, unlike the previous cases discussed in the literature, where the global ground and/or thermal states describing the overall system were studied, we will consider partial quantum states of certain subsystems instead. We will consider a particular physical system, a conventional BCS superconductor with an inserted magnetic impurity, and investigate the behavior of the zero-temperature fidelity between the partial mixed states, corresponding to subsystems chosen, rather than between the global pure ground states, corresponding to the overall system. We will show that in the model considered, the partial state fidelity can mark the point of the quantum phase transition in the same manner as in the case of previously studied global fidelities corresponding to the overall ground states. We will also establish a connection between the order parameter and a so-called optimal observable given by the fidelity (see the discussion from the next section), identifying the on-site magnetization as the order parameter for the phase transition considered. Finally, we

 ©2008 The American Physical Society

will study the behavior of the Uhlmann mixed state geometric phase and, using partial state fidelities, discuss the amount of total (both classical and quantum) correlations in the vicinity of the point of the quantum phase transition.

The paper is organized as follows. In the next section, we introduce the formal notion of fidelity, presenting in brief its connection with the study of phase transitions (both quantum and thermal). In particular, we address the connection between the order parameter and an optimal observable determined by the fidelity. In Sec. III, we give the Hamiltonian of the system, discussing some basic features of the model considered, as well as the analytic techniques used. Section IV contains the main part of this paper. There, we present the basic numerical techniques used in computing the partial one- and two-site states, together with the results for the corresponding partial one- and two-site fidelities. We show that the point of the phase transition is clearly marked by sudden drops of suitably chosen partial states fidelities. We also show that this particular choice of fidelities (i.e., the choice of subsystems considered) enables us to identify the very order parameter as well—the on-site magnetization. Finally, in Sec. V we summarize conclusions and discuss some future directions of research.

II. FIDELITY AND PHASE TRANSITIONS

Given two physical systems a and b, their quantum states, being either pure or mixed, are given by density operators $\hat{\rho}_a$ and $\hat{\rho}_b$. Since according to quantum mechanics it is in general not possible to unambiguously distinguish with probability one between two given states [1,23], various measures of quantum state distinguishability have been introduced (for an overview, see [23]). One of the most widely used measures is (quantum) fidelity [4], given by the expression

$$F(\hat{\rho}_a,\hat{\rho}_b) = \mathrm{Tr}\sqrt{\sqrt{\hat{\rho}_a}\hat{\rho}_b\sqrt{\hat{\rho}_a}}. \tag{1}$$

The numerical value of the fidelity ranges from $F=0$, in the case of fully distinguishable states [38], to $F=1$, when the two quantum states are completely indistinguishable from each other (i.e., they are identical). As it happened with most other quantum state distinguishability measures, the introduction of the quantum fidelity F was also motivated by the analogous classical quantity. Classical fidelity F_c, a measure of distinguishability between two probability distributions $\{p_a(i)\}$ and $\{p_b(i)\}$, is defined as $F_c(p_a,p_b)=\Sigma_i\sqrt{p_a(i)p_b(i)}$. Measuring an observable $\hat{A}$ in states $\hat{\rho}_a$ and $\hat{\rho}_b$, one obtains probability distributions $\{p_a^{\hat{A}}(i)\}$ and $\{p_b^{\hat{A}}(i)\}$, respectively, whose mutual distinguishability is given by the classical fidelity $F_c(p_a^{\hat{A}},p_b^{\hat{A}})$. It can be proven (see [23]) that for any observable $\hat{A}$, the inequality $F(\hat{\rho}_a,\hat{\rho}_b)\leq F_c(p_a^{\hat{A}},p_b^{\hat{A}})$ holds, and there always exists an optimal observable $\hat{A}_{op}$ for which the equality is achieved:

$$F(\hat{\rho}_a,\hat{\rho}_b) = \mathrm{Tr}\sqrt{\sqrt{\hat{\rho}_a}\hat{\rho}_b\sqrt{\hat{\rho}_a}} = \sum_i \sqrt{p_a^{\hat{A}_{op}}(i)p_b^{\hat{A}_{op}}(i)}$$
$$= F_c(p_a^{\hat{A}_{op}},p_b^{\hat{A}_{op}}). \tag{2}$$

In other words, quantum fidelity $F(\hat{\rho}_a,\hat{\rho}_b)$ gives the optimal value for the distinguishability between two quantum states, when compared through probability distributions $\{p_a^{\hat{A}}(i)\}$ and $\{p_b^{\hat{A}}(i)\}$ (note though that it does not give an optimal observable $\hat{A}_{op}$). Using the modulus of an operator $\hat{R}$, $|\hat{R}|=(\hat{R}\hat{R}^{\dagger})^{1/2}$, one obtains an alternative expression for the fidelity [1],

$$F(\hat{\rho}_a,\hat{\rho}_b) = \mathrm{Tr}|\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}| = \mathrm{Tr}\sqrt{\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}(\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b})^{\dagger}}$$
$$= \mathrm{Tr}\sqrt{\sqrt{\hat{\rho}_a}\hat{\rho}_b\sqrt{\hat{\rho}_a}}. \tag{3}$$

The fidelity approach to quantum phase transitions is based on the use of the above quantity (1) in distinguishing macroscopic quantum states of many-body quantum systems. Consider a system given by the set of Hamiltonians $\hat{H}(q)$, where q is a parameter representing a set of interaction coupling constants, which at $T=0$ is in a pure ground state $|\Phi\rangle=|\Phi(q)\rangle$. Then, the sudden drop of the fidelity (1) $F(|\Phi(q)\rangle\langle\Phi(q)|,|\Phi(\tilde{q})\rangle\langle\Phi(\tilde{q})|)=|\langle\Phi(q)|\Phi(\tilde{q})\rangle|$ between two ground states $|\Phi(q)\rangle$ and $|\Phi(\tilde{q})\rangle$ obtained for two close parameter points q and $\tilde{q}=q+\delta q$, with δq being small, can indicate the points of quantum phase transitions. The basic logic behind this idea is strikingly simple—two quantum states defining different macroscopic phases are expected to have enhanced distinguishability that would quantitatively exceed that taken between states from the same phase.

Yet, it is precisely the simplicity where one of the advantages of this approach lies: being rooted on a basic notion such as state distinguishability, the fidelity study provides a general approach to phase transitions. From Eq. (2), we see that a sudden drop of the fidelity $F(\hat{\rho}_a,\hat{\rho}_b)$ corresponds to a nonanalyticity in the classical fidelity $F_c(p_a^{\hat{A}_{op}},p_b^{\hat{A}_{op}})$, given by an optimal observable $\hat{A}_{op}$. In other words, there always exists an observable such that, when measured in the states $\hat{\rho}_a$ and $\hat{\rho}_b$, it generates probability distributions whose (classical) distinguishability is enhanced. This suggests a natural choice for an order parameter, given by an optimal observable $\hat{A}_{op}$. Note that in this approach, to study phase transitions one does not need to know the order parameter in advance. Equation (2) ensures that the behavior of the fidelity will be directly determined by the existence of the order parameter, given by an optimal observable $\hat{A}_{op}$.

The relation between the fidelity and the thermodynamic response functions provides another connection between the fidelity drop and phase transitions. For the case of quantum phase transitions described by the Landau-Ginzburg-Wilson (LGW) theory, where the overall Hamiltonian of the system

has the simplified form $\hat{H}=\hat{H}_0-h\hat{S}$, with $\hat{S}$ being the symmetry-breaking term and $h=h(q)$, the following relation holds (see, for example, [17]):

$$F|_{h=\Delta h} \simeq e^{-1/2\chi\Delta h^2}, \tag{4}$$

where $\chi=\int_0^\infty \tau[\langle\hat{S}(\tau)\hat{S}(0)\rangle-\langle\hat{S}(0)\rangle^2]d\tau$, with $\hat{S}(\tau)=e^{\tau\hat{S}}\hat{S}e^{-\tau\hat{S}}$, is the corresponding zero-field dynamical susceptibility, associated with the symmetry-breaking field, and $\Delta h=h_a-h_b$. We see that for this broad class of systems described by the LGW theory, the nonanalyticity of the susceptibility, a signature of the quantum phase transitions, is directly reflected in the sudden drop of the fidelity. The analogous relation holds for thermal phase transitions of the LGW type, for the systems for which Hamiltonians $\hat{H}(q_a)$ and $\hat{H}(q_b)$ (at the parameter points q_a and q_b) commute with each other (see, for example, [20]).

As already mentioned in the introduction, in this paper we will focus on partial, rather than global states of the system, and explore the behavior of the corresponding partial state fidelities. If the overall system $\mathcal{S}$ is divided into two parts A and B, then we will study the zero-temperature mixed state fidelity $F[\hat{\rho}_A(q),\hat{\rho}_A(\tilde{q})]$, where partial states are given by mixed states $\hat{\rho}_A(q)=\mathrm{Tr}_B|\Phi(q)\rangle\langle\Phi(q)|$, and analogously for $\hat{\rho}_B(\tilde{q})$. We will show that for the system considered, a conventional BCS superconductor with an inserted magnetic impurity, a sudden drop of the above mixed state fidelity can define the point at which the (first-order) quantum phase transition occurs, even when applied to the smallest subsystems of the overall system. Also, we will identify the order parameter, the on-site magnetization, as an optimal observable given by the one-site partial fidelity. Note that, since in this particular case the transition is of the first order, occurring already for a finite size of the system, there is no need for the analysis of the scaling behavior in this model.

III. MODEL

In this work we consider the quantum phase transition induced by a magnetic impurity inserted in a conventional superconductor [24–31]. Consider a classical spin immersed in a two-dimensional s-wave conventional superconductor. We use a lattice description of the system. In the center of the system, $i=l_c=(x_c,y_c)$, we place a classical spin parametrized by

$$\frac{\vec{S}}{S}=\cos\varphi\vec{e}_x+\sin\varphi\vec{e}_z. \tag{5}$$

The Hamiltonian of the system is given by

$$\hat{H}=-\sum_{\langle ij\rangle\sigma} t_{ij}\hat{c}^\dagger_{i\sigma}\hat{c}_{j\sigma}-\varepsilon_F\sum_{i\sigma}\hat{c}^\dagger_{i\sigma}\hat{c}_{i\sigma}+\sum_i(\Delta_i\hat{c}^\dagger_{i\uparrow}\hat{c}^\dagger_{i\downarrow}+\Delta_i^*\hat{c}_{i\downarrow}\hat{c}_{i\uparrow})$$
$$-\sum_{\sigma\sigma'}J(\cos\varphi\hat{c}^\dagger_{l_c\sigma}\sigma^x_{\sigma\sigma'}\hat{c}_{l_c\sigma'}+\sin\varphi\hat{c}^\dagger_{l_c\sigma}\sigma^z_{\sigma\sigma'}\hat{c}_{l_c\sigma'}), \tag{6}$$

where the first term describes the hopping of electrons between different sites on the lattice, ε_F is the chemical potential (equal to the Fermi energy at $T=0$), the third term is the superconducting s pairing with the site-dependent order parameter Δ_i, and the last term with $J>0$ is the exchange interaction of an electron at site $i=l_c$ with the magnetic impurity. The hopping matrix is given by $t_{ij}=t\delta_{j,i+\delta}$, where δ is a vector to a nearest-neighbor site. Note that both the indices l and $i,j\in\{1,2,\ldots,N\}$ specify sites on a two-dimensional system (N is the number of sites). We will take energy units in terms of t ($t=1$), and $\varepsilon_F=-1$.

Since the impurity spin acts as a local magnetic field the electronic spin density will align along the local spin. For small values of the coupling there is a negative spin density around the impurity site. At the impurity site it is positive, as expected. For larger couplings the spin density in the vicinity of the impurity site is positive. At small couplings the many-body system screens the effect induced by the impurity inducing fluctuations that compensate the effect of the local field in a way that the overall magnetization vanishes. However, for a strong enough coupling the many-body system becomes magnetized in a discontinuous fashion. One interpretation is that if J is strong enough the impurity breaks a Cooper pair and captures one of the electrons, leaving the other electron unpaired, and thus the overall electronic system becomes polarized. The impurity induces a pair of bound states inside the superconducting energy gap, one at positive energy (with respect to the chemical potential), and another at a symmetric negative energy. Even though the spectrum is symmetric the spectral weights of the two energy levels are not the same and their spin content is also distinct. The analysis of the local density of states (LDOS) [28,32] shows that for small coupling the lowest positive energy level has only a contribution from spin $\uparrow$ and the first level with negative energy (symmetric to the other level) has only a contribution from spin $\downarrow$. The magnitude of the spectral weight at the impurity site is different for the two states. Considering a higher value for the coupling one finds that the levels inside the gap approach the Fermi level. There is a critical value of the coupling for which the two levels cross in a discontinuous way such that it coincides with the emergence of a finite overall magnetization. After the level crossing occurs, the nature of the states changes. The positive energy bound state has now only a contribution from the spin $\downarrow$ component and vice versa, the first negative energy state has only contribution from the spin component $\uparrow$—hence as the level crossing takes place the spin content changes.

The diagonalization of this Hamiltonian is performed using the Bogoliubov-Valatin transformation in the form

$$\hat{c}_{i\uparrow}=\sum_n[u_n(i,\uparrow)\hat{\gamma}_n-v_n^*(i,\uparrow)\hat{\gamma}_n^\dagger],$$

$$\hat{c}_{i\downarrow}=\sum_n[u_n(i,\downarrow)\hat{\gamma}_n+v_n^*(i,\downarrow)\hat{\gamma}_n^\dagger]. \tag{7}$$

Here n is a complete set of states, u_n and v_n are related to the eigenfunctions of the Hamiltonian (6), and the new fermionic operators $\hat{\gamma}_n$ are the quasiparticle operators. These are chosen such that in terms of the new operators

$$\hat{H} = E_g + \sum_n \epsilon_n \hat{\gamma}_n^\dagger \hat{\gamma}_n, \tag{8}$$

where E_g is the ground-state energy and ϵ_n are the excitation energies. As a consequence,

$$[\hat{H}, \hat{\gamma}_n] = -\epsilon_n \hat{\gamma}_n,$$

$$[\hat{H}, \hat{\gamma}_n^\dagger] = \epsilon_n \hat{\gamma}_n^\dagger. \tag{9}$$

The coefficients $u_n(i,\sigma)$, $v_n(i,\sigma)$ can be obtained solving the Bogoliubov–de Gennes (BdG) equations [33]. Defining the vector

$$\psi_n(i) = \begin{pmatrix} u_n(i,\uparrow) \\ v_n(i,\downarrow) \\ u_n(i,\downarrow) \\ v_n(i,\uparrow) \end{pmatrix}, \tag{10}$$

the BdG equations can be written as

$$\mathcal{H}\psi_n(i) = \epsilon_n \psi_n(i), \tag{11}$$

where the $\mathcal{H}$ matrix at site i is given by

$$\mathcal{H} = \begin{pmatrix} -h - \varepsilon_F - J\sin\varphi & \Delta_i & -J\cos\varphi & 0 \\ \Delta_i^* & h + \varepsilon_F - J\sin\varphi & 0 & -J\cos\varphi \\ -J\cos\varphi & 0 & -h - \varepsilon_F + J\sin\varphi & \Delta_i \\ 0 & -J\cos\varphi & \Delta_i^* & h + \varepsilon_F + J\sin\varphi \end{pmatrix}, \tag{12}$$

where $h = t\hat{s}_\delta$ with $\hat{s}_\delta f(i) = f(i+\delta)$. The solution of these equations gives both energy eigenvalues and eigenstates. The problem involves the diagonalization of a $(4N)\times(4N)$ matrix. The solution of the BdG equations is performed self-consistently imposing at each iteration that

$$\Delta_i = \frac{g}{2}[\langle \hat{c}_{i\uparrow}\hat{c}_{i\downarrow}\rangle - \langle \hat{c}_{i\downarrow}\hat{c}_{i\uparrow}\rangle], \tag{13}$$

where g is the effective attractive interaction between the electrons. Using the canonical transformation this can be written as

$$\Delta_i = -g \sum_{n,0<\epsilon_n<\hbar\omega_D} \left(f_n(\epsilon_n) - \frac{1}{2}\right)[u_n(i,\uparrow)v_n^*(i,\downarrow) + u_n(i,\downarrow)v_n^*(i,\uparrow)] = \frac{g}{2} \sum_{n,0<\epsilon_n<\hbar\omega_D} \tanh\frac{\beta\epsilon_n}{2}[u_n(i,\uparrow)v_n^*(i,\downarrow) + u_n(i,\downarrow)v_n^*(i,\uparrow)], \tag{14}$$

where ω_D is the Debye frequency, and $f_n(\epsilon_n)$ is the Fermi function defined as usual as

$$f_n(\epsilon_n) = \frac{1}{e^{\beta\epsilon_n} + 1}, \tag{15}$$

where $\beta = 1/(k_B T)$, with k_B being the Boltzmann constant and T the absolute temperature. We note that the Bogoliubov–de Gennes equations are invariant under the substitutions $\epsilon_n \to -\epsilon_n$, $u(\uparrow) \to v(\uparrow)$, $v(\uparrow) \to u(\uparrow)$, $v(\downarrow) \to -u(\downarrow)$, $u(\downarrow) \to -v(\downarrow)$.

IV. RESULTS

As we have already noted, in this paper we focus on the fidelity between the partial states rather than the overall ground states, as it was the case in the previous studies [5–7,9,11–20]. At $T=0$, the system is in the pure ground state $|\Phi\rangle$ and has a density operator $\hat{\rho} = |\Phi\rangle\langle\Phi|$. Thus, if we divide the whole system in two subsystems, say A and B, then the partial mixed state, given by the reduced density operator $\hat{\rho}_A$ for the subsystem A, is defined as

$$\hat{\rho}_A = \mathrm{Tr}_B\, \hat{\rho}, \tag{16}$$

where $\mathrm{Tr}_B(\cdots)$ represents the partial trace evaluated over the Hilbert space $\mathcal{H}_B$ of the subsystem B. In particular, we will calculate the matrix elements of one-site and two-site partial mixed states, given in general for different sites i and (i,j), respectively, and different values of the coupling constant J. In order to infer the point of the quantum phase transition, we will calculate the fidelity between two one-site and two two-site states, given for two close parameters J and $J+\delta J$. Further, to better understand the influence of the quantum phase transition point on the structure of partial states (and therefore, to a certain extent, the overall state as well), we examine the fidelity between two different partial states, given for the same values of the coupling constant J, but different sites, for both cases of one- and two-site states.

In many-body systems a description in terms of quantum pure states and wave functions as their representatives in a certain basis (e.g., position, momentum, energy, etc.) is quite involved and the second quantization is the natural way to perform any calculation. The mixed states are however easily defined in terms of pure states and their matrix elements in

some basis, and this approach is indeed not difficult to implement using Fock states. The matrix elements of the density matrix are simply defined in terms of correlation functions of the whole system. For instance, in the case of the single-site partial mixed states it can be shown that, using local basis states $\mathcal{B}=\{|0\rangle,|\uparrow\downarrow\rangle|\uparrow\rangle,|\downarrow\rangle\}$, which denote the four possible states—unoccupied, double occupied, single occupied with an electron with spin up, and single occupied with an electron with spin down, respectively—the corresponding density matrix reads as

$$\rho_i = \begin{pmatrix} \langle(1-\hat{n}_\uparrow)(1-\hat{n}_\downarrow)\rangle & \langle\hat{c}_\uparrow^\dagger\hat{c}_\downarrow^\dagger\rangle & 0 & 0 \\ \langle\hat{c}_\downarrow\hat{c}_\uparrow\rangle & \langle\hat{n}_\uparrow\hat{n}_\downarrow\rangle & 0 & 0 \\ 0 & 0 & \langle\hat{n}_\uparrow(1-\hat{n}_\downarrow)\rangle & \langle\hat{c}_\downarrow^\dagger\hat{c}_\uparrow\rangle \\ 0 & 0 & \langle\hat{c}_\uparrow^\dagger\hat{c}_\downarrow\rangle & \langle(1-\hat{n}_\uparrow)\hat{n}_\downarrow\rangle \end{pmatrix}_i, \tag{17}$$

where the index i denotes the site. Note that it can be written as an orthogonal sum $\rho_i=\rho_i^c\oplus\rho_i^s$, where ρ_i^c and ρ_i^s are the matrix representations of the partial states corresponding to the charge and the spin degrees of freedom, respectively, which therefore are decoupled. The charge part couples the empty and doubly occupied states and the spin part couples the single occupied states (the two spin orientations). The diagonal terms of the matrix describe the number of empty sites, the number of doubly occupied sites, the number of spin-up sites and the number of spin-down sites, respectively. The sum of the diagonal terms is equal to 1 due to normalization. The matrix is easily diagonalized and the fidelity (1) between two different one-site states obtained straightforwardly.

The correlation functions are easily solved using the representation of the electron operators in terms of the BdG quasiparticle operators. Specifically, we can write the single-site density matrix as

$$\hat{\rho} = \sum_{m,n} \rho_{mn}|m\rangle\langle n|, \tag{18}$$

where $|m\rangle,|n\rangle\in\mathcal{B}$ are the four basis states described above. Consider now an operator $\hat{O}$ defined in terms of creation and annihilation operators of the electrons, and evaluate

$$\mathrm{Tr}\{\hat{O}\hat{\rho}\} = \sum_{m,n} \langle n|\hat{O}|m\rangle\rho_{mn}. \tag{19}$$

To determine the matrix element ρ_{mn} of the density matrix it is enough, by inspection, to find which operator $\hat{O}$ is such that only the matrix element $\langle n|\hat{O}|m\rangle$ is nonzero. This will tell us which correlation function gives which matrix element of the density matrix. Alternatively, one can express the operator $|m\rangle\langle n|$ in terms of the creation and annihilation operators. The results for ρ_A are given by Eq. (17) above.

These matrix elements are defined in the subspace of the states of one site. However, since we have integrated out all the other sites, we may now replace the matrix elements by the matrix elements over the entire system (direct product with the states of the remaining $N-1$ sites).

Computing the fidelity between the two-site partial states is more involved. Given two sites, i and j, the two-site basis states are given by the direct product of the one-site basis sets that were defined earlier, and the reduced density matrix is now a square matrix of order 16. However, exactly one-half of the 256 matrix elements are null—the nonzero ones being those that correspond to either even-even or odd-odd states, where the parity reflects the number of electrons in the sites i and j, respectively. Hence those two sets of states decouple and the reduced density matrix can be written as a block diagonal matrix where each submatrix is a square matrix or order 8. Still, the calculation is rather complex and for that reason we used the computer program FORM [34] to obtain symbolic expressions for the matrix elements involved. Note that one of the correlation functions involves the double occupancies at the two sites, and thus one should deal with products of up to eight electron operators (where each electron operator is given by a sum of two quasiparticle operators); also, the signs that follow from the anticommutation relations of fermionic operators must be taken into account.

A. Fidelity between single-site states

In the following, we present the numerical results for the fidelity between one-site partial states, for a two-dimensional lattice of size 15×15 sites. For computational simplicity, we perform our calculations for this system size. Since the effect of the impurity is short ranged, changing the size of the system only affects slightly the results. First, we evaluate the fidelity (1) between two one-site states given for two sites i and j and two close values J and $J+\delta J$ of the exchange parameter, with $\delta J=0.05$ [39], such that $\hat{\rho}_a=\hat{\rho}(J;i)$ and $\hat{\rho}_b=\hat{\rho}(J+\delta J;j)$.

For $i=j=l_c$, when the electrons are located at the site of the magnetic impurity, the results for the one-site fidelity $F_1(\hat{\rho}_a,\hat{\rho}_b)$, as a function of J, are presented in Fig. 1. We observe a clear drop, from the otherwise common value $F_1(\hat{\rho}_a,\hat{\rho}_b)\simeq 1$, for the critical value $J_0\simeq 1.9$ [32] of the parameter J. As we move one of the electrons from the impurity, the drop of the fidelity F_1 around the point of the quantum phase transition becomes much smaller: For the first

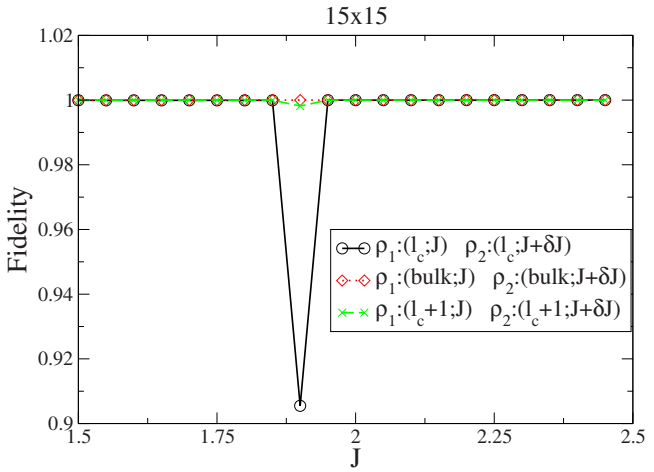


FIG. 1. (Color online) One-site fidelity between two states taken for two close parameter points J and $J+\delta J$ ($\delta J=0.05$), for the site at the impurity (black line), a nearest neighbor (green line), and a site located in the bulk (red line). Note the sudden drop of the fidelity for the impurity around the point of the quantum phase transition.

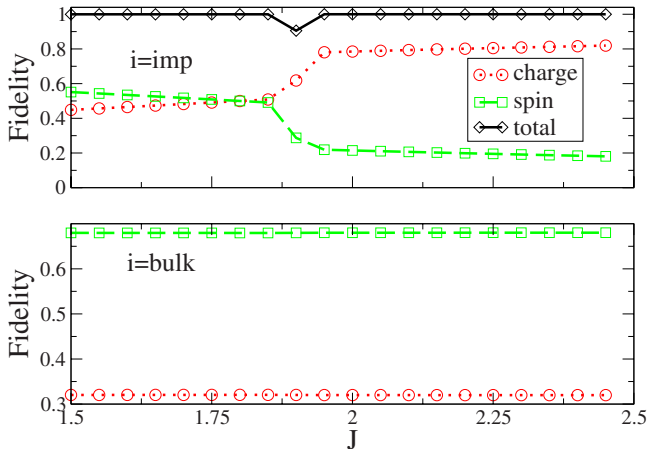


FIG. 2. (Color online) The charge and spin contributions of the one-site fidelity as a function of the parameter J, with $\delta J=0.05$, for the site at the impurity (upper figure) and a site located in the bulk (lower figure). Note the discontinuity for both spin and charge contributions around the point of the quantum phase transition, for the site located at the impurity.

neighbors of the impurity, it is still visible, while already for the second neighbors and the far away bulk sites, it becomes negligible. We see that even for the smallest subsystems—one-site spins in the lattice—the fidelity can be dramatically influenced by the structure of their states, as long as they are close to the impurity. This is intuitively easy to accept since it was shown in the previous work [32,35], through investigation of various relevant physical quantities, such as the gap, the electron density or entanglement (quantum correlations), and the mutual information (total correlations—classical and quantum) that the impurity affects one-site properties of the system mainly in its vicinity (although it was also shown that for some two-site entanglement measures the effects persisted for larger distances).

A more detailed picture of the fidelity between the one-site states is presented in Fig. 2, where the two distinct contributions to the overall change of partial states is presented. The charge and the spin parts ρ_i^c and ρ_i^s of the one-particle state given by $\rho_i=\rho_i^c\oplus\rho_i^s$ [see Eq. (17) and the discussion below] lead to two different contributions to the state distinguishability, and we see that they both exhibit a discontinuity in the vicinity of the point of the quantum phase transition.

Further, we have evaluated the Uhlmann mixed state geometric phase [22], a generalization of the Berry geometric phase to mixed quantum states. While the connection between Berry phases, pure state fidelity and quantum phase transitions, was studied before [15,16,21], the Uhlmann mixed state geometric phase was recently introduced [20] to analyze the structural change of the system eigenvectors, with respect to the change of the relevant parameter driving the system Hamiltonian (i.e., J in our case). The Uhlmann geometric phase ϕ_g (and the corresponding holonomy) is determined by the unitary operator $\hat{V}_{ab}$, given by the polar decomposition (see, for example, [1]) $\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}=|\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}|\hat{V}_{ab}$, where $|\hat{R}|=(\hat{R}\hat{R}^\dagger)^{1/2}$ represents the modulus of the operator $\hat{R}$ (for details, see [36]). Since the fidelity can be written as $F(\hat{\rho}_a,\hat{\rho}_b)=\mathrm{Tr}|\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b}|$, see Eq. (3), a nonzero difference $H(\hat{\rho}_a,\hat{\rho}_b)-F(\hat{\rho}_a,\hat{\rho}_b)$, where $H(\hat{\rho}_a,\hat{\rho}_b)=\mathrm{Tr}(\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b})$, would imply $\hat{V}_{ab}\neq\hat{I}_{ab}$ and therefore the emergence of a nontrivial Uhlmann geometric phase $\phi_g\neq 0$. On the other hand, for the case of mutually commuting Hamiltonians the system eigenvectors do not change with the driving parameter and we have $F(\hat{\rho}_a,\hat{\rho}_b)=\mathrm{Tr}(\sqrt{\hat{\rho}_a}\sqrt{\hat{\rho}_b})$. In other words, $F(\hat{\rho}_a,\hat{\rho}_b)=H(\hat{\rho}_a,\hat{\rho}_b)$ and the Uhlmann geometric phase becomes trivial, i.e., $\phi_g=0$.

For the one-site states, we find that the Uhlmann geometric phase is always trivial for every site neighboring the impurity or in the bulk: The one-site eigenvectors do not contribute significantly to the change of the overall one-site partial states at the point of the quantum phase transition. Still, for the impurity site itself, around the point of the phase transition J_0 we observe a small deviation from zero for the difference H_1-F_1, which indicates a slight "rotation" of the electron's eigenbasis. This result can be qualitatively explained noticing that in the model we are considering the third term in the Hamiltonian (6), giving the superconducting features, and the last term in the Hamiltonian, giving the coupling between the impurity and the lattice electrons, commute with each other [40]. Moreover, for a fixed value of the gap Δ_i, changing the impurity-lattice coupling constant J, without changing the impurity spin orientation angle φ, will result in no change of the overall Hamiltonian's eigenbasis—only the eigenvalues are affected in this case and the system exhibits a first-order quantum phase transition as a consequence of the explicit (not avoided) level crossing. Thus, one might expect that also the partial subsystem states would feature little change of local eigenbasis, even in the vicinity of the point of the quantum phase transition. This is precisely the case in the model considered, apart for the gap Δ_{l_c} given for the electron located at the impurity (see [32]): For every site i, the gap Δ_i is a slowly changing function of J, while for the impurity Δ_{l_c} exhibits a π shift and becomes negative. Such a behavior of Δ_{l_c} does affect the difference H_1-F_1, but the effect is an order of magnitude smaller than for the fidel-

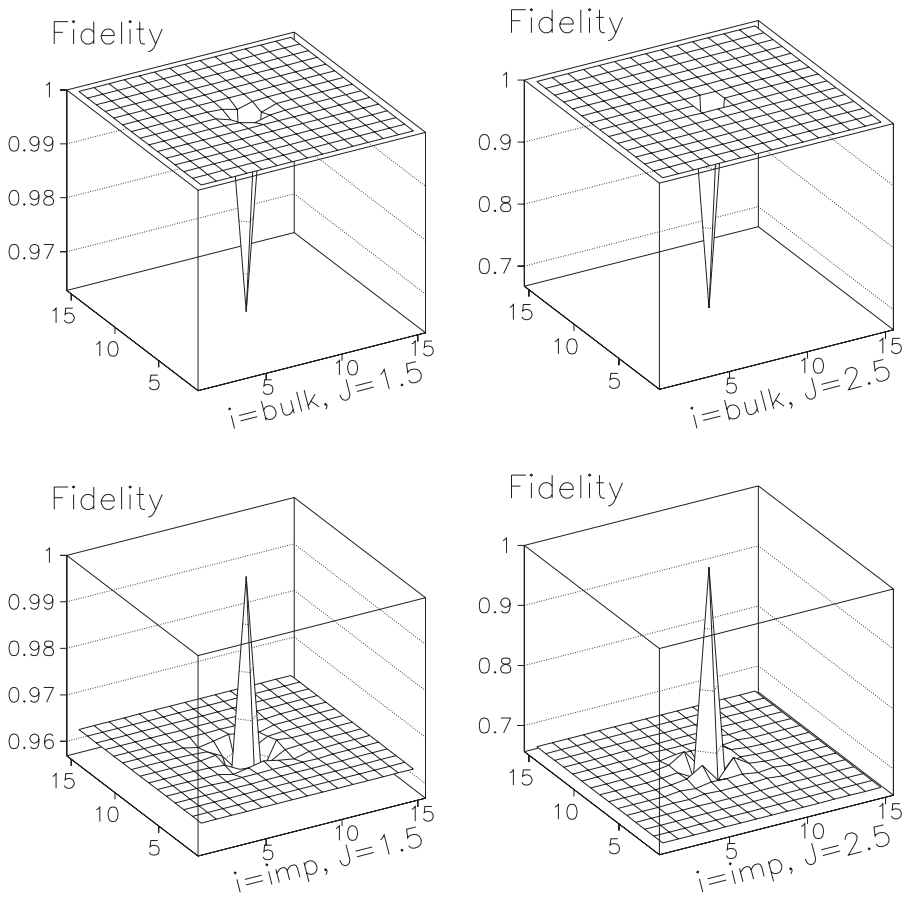


FIG. 3. One-site fidelity, given for fixed values of the parameter J, between any site and a site in the bulk (upper figures); a site in the impurity (lower figures). Left-hand figures are taken for $J^-=1.5<J_0$, figures on the right-hand side for $J^+=2.5>J_0$, where $J_0=1.9$ is the point of the quantum phase transition. Note the drop of the fidelity between the impurity and the bulk state, from $F_1^-\simeq 0.96$ for $J^-=1.5<J_0$ to $F_1^+\simeq 0.65$ for $J^+=2.5>J_0$—another signature of the quantum phase transition.

ity. Further, it affects the impurity site only and thus does not lead to the superconductor-normal metal type of phase transition.

Finally, to further investigate the structure of one-site partial states, we have evaluated the fidelity between the states for two different sites, and the same value of the parameter J. The results are presented in Fig. 3. Again, we see that apart from the state of the site l_c located at the impurity, and those at its vicinity (first neighbors), all the other one-site partial bulk states are almost the same, for a given common value of the parameter J. In Fig. 3 we have presented the results for two distinct values of the parameter, $J^-=1.5<J_0$ and $J^+=2.5>J_0$, but the qualitative picture stays the same for every $J^-<J_0$ and $J^+>J_0$: The numerical value for the fidelity between the bulk state (say, $\hat{\rho}_1$) and the impurity state (say, $\hat{\rho}_2$) is significantly greater for J^- than for J^+, in other words, $F_1^-(\hat{\rho}_a,\hat{\rho}_b)>F_1^+(\hat{\rho}_a,\hat{\rho}_b)$, thus revealing the point of the quantum phase transition.

Note that when $i=l_c$ and $J<J_0$, besides the peak at the impurity site, in its vicinity the fidelity F_1 is smaller than the background value, while for $J>J_0$ it is larger. This is directly connected to the existence of oscillations of the magnetization around the impurity site (see [32]). Within the bounds set by quantum mechanics, two quantum states can be distinguished by measuring different observables and comparing the outcomes obtained. Magnetization is one of them, and in fact we see that it is optimal, or at least nearly so: For $J<J_0$, it is easier to distinguish the impurity site from a site in its immediate vicinity than from other, more distant sites, simply because the difference in the expected magnetization between two sites is bigger in the case of impurity neighborhood than in other cases. Therefore, in this case the fidelity has a drop in the neighborhood of the impurity, i.e., the distinguishability is larger there than elsewhere. A similar analysis holds for $J>J_0$.

B. Fidelity between two-site states

We have also studied the behavior of the fidelity taken between two-site partial states. We considered two distinct cases. First, we calculated the fidelity between two-site states for which one of the sites i,j is at the impurity, say $i=l_c$, and the other is in its vicinity, $j=l_c+\delta l$ ($\delta l=1,2,\ldots,6$), for two close parameter values J and $J+\delta J$, with $\delta J=0.05$. In Fig. 4 we present the results for the two-site fidelity $F_2(\hat{\rho}_a,\hat{\rho}_b)$, with $\hat{\rho}_a=\hat{\rho}(J;l_c,l_c+\delta l)$ and $\hat{\rho}_b=\hat{\rho}(J+\delta J;l_c,l_c+\delta l)$, where for clarity we plotted only the $\delta l=1,2,3$ cases, as the other curves almost overlap the results for $\delta l=3$. Again, as in the previous case of the fidelity between the one-site states, here as well the sudden drop of the fidelity from its common value

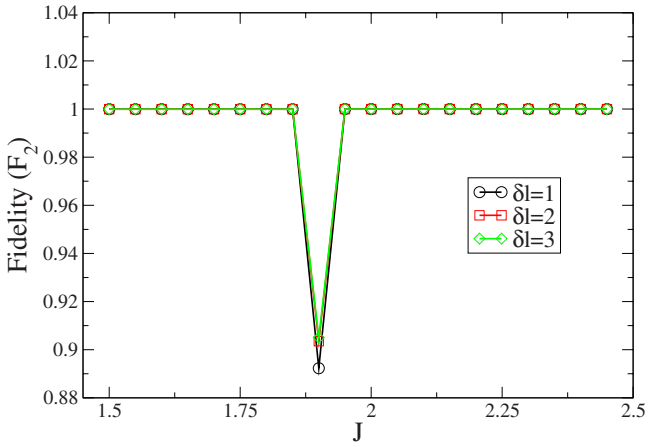


FIG. 4. (Color online) Two-site fidelity between two states taken for two close parameter points J and $J+\delta J$ ($\delta J=0.05$), when one site is at the impurity and the other site in the neighborhood of it. Note the sudden drop of the fidelity around the point of the quantum phase transition, for the case of impurity+neighboring site.

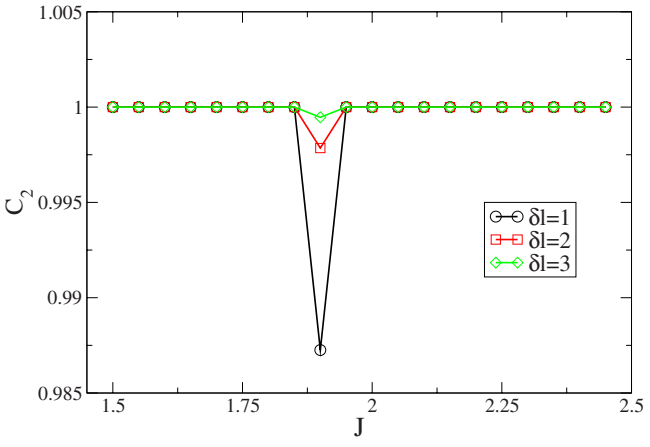


FIG. 5. (Color online) The C_2 function, defined by Eq. (20), between two states taken for two close parameter points J and $J+\delta J$ ($\delta J=0.05$), when one site is at the impurity and the other site in the neighborhood of it. Note its sudden drop around the point of the quantum phase transition, as well as the decrease of its deviation from 1 as the other site gets further from the impurity.

$F_2(\hat{\rho}_a,\hat{\rho}_b)\simeq 1$ can be observed precisely in the vicinity of the critical point $J_0=1.9$. We see it as a signature of both the change in the impurity one-site partial states, as well as the change of the total correlations between the impurity and the remaining sites. From our study of the one-site fidelity, we have seen that the bulk partial states are all approximately the same, irrespectively of the site and the parameter value, while only the impurity and first-neighbor one-site states exhibit a visible dramatic change around the point of the quantum phase transition J_0.

On the other hand, for every value of J we can consider the quantity

$$C_2(J,J+\delta J;l_c,l_c+\delta l)\equiv\frac{F_2[\hat{\rho}(J;l_c,l_c+\delta l),\hat{\rho}(J+\delta J;l_c,l_c+\delta l)]}{F_1[\hat{\rho}(J;l_c),\hat{\rho}(J+\delta J;l_c)]F_1[\hat{\rho}(J;l_c+\delta l),\hat{\rho}(J+\delta J;l_c+\delta l)]},\tag{20}$$

which can be seen as an indicator of total correlations, as in the case of uncorrelated two-site composite systems the above quotient reduces identically to 1. Therefore, the two-site composite system is necessarily correlated whenever $C_2\neq 1$. Note that C_2 is not the measure of correlations—the two-site composite system can still exhibit nonzero correlations between its one-site subsystems even for $C_2=1$. This is clearly seen in Fig. 5, where $C_2=1$ for all values of J except for $J=J_0$, although is was shown in [35] that the considered two-site systems are correlated for every value of J. Yet, precisely around the point J_0 of the phase transition, the quantity C_2 exhibits an abrupt change, thus showing an enhanced change in the amount of the total correlations between the one-site subsystems of an overall two-site system, see Fig. 5. We also see that its numerical value around $J=J_0$ approaches 1 as the step δl increases—i.e., the change of total correlations decreases as the distance between the two sites increases. This behavior was already observed before in the study of total and quantum correlations (entanglement) in this model [35]. In connection to this, we note here that the fidelity induces various distance measures on the set of quantum states (for an overview see, for example, [23] and references therein), which can further induce fidelity-based multiparticle entanglement measures, based on the quantum state distinguishability, similarly to the case of the well-known relative entropy of entanglement measures [37].

In the case of two-site subsystems with both one-site electrons located in the bulk, the fidelity taken between states obtained for two close parameter values J and $J+\delta J$, is essentially equal to 1 (with $\delta J=0.05$). Analogously to the case of the fidelity between one-site partial states, here as well we see that the impurity does not affect considerably the partial states of the bulk of the lattice, even around the point of the quantum phase transition. Yet, we have observed a small deviation around J_0 from the otherwise common value $F_2\simeq 1$ which indicates that for larger subsystems (when considering the fidelities F_3, F_4, etc., of the three-electron and four-electron subsystems, respectively), the impurity might start to increasingly influence the collective behavior of the bulk.

Finally, as in the previous case, we have evaluated the Uhlmann mixed state geometric phase ϕ_g, again obtaining the trivial value $\phi_g\simeq 0$ everywhere, apart for the case when one electron site is located at the impurity: The structure of

the two-site partial states' eigenvectors is not considerably affected by the quantum phase transition.

V. CONCLUSION

The study of the fidelity provides a general approach to phase transitions. It is based on a simple notion of state distinguishability, quantified by the fidelity, and does not require any *a priori* knowledge of standard properties previously used to describe phase transitions, such as order parameters, symmetry orders, behavior of particular correlation functions, entanglement measures, etc. Nevertheless, its behavior is, via Eq. (2), directly determined by the behavior of the optimal observable $\hat{A}_{op}$, which can be used as an order parameter (in the case considered in this paper, it was the magnetization, discussed in detail in the preceding section). The fidelity approach to both quantum and thermal phase transitions was proven to be correct for numerous particular, as well as for broad classes of physical systems. It has also shown to be useful in connecting different theoretical concepts related to phase transitions, such as the behavior of the fidelity and the emergence of nontrivial geometric (Berry and Uhlmann) phases along the regions of criticality (see [15,16,20]), or the relation between the fidelity and the Loschmidt echo of the environment, a quantity that determines the decoherence rates of quantum systems interacting with the bath (see [5]).

In this paper we have applied the concept of fidelity to approach quantum phase transitions, taking as the physical system a conventional two-dimensional superconductor with a classical magnetic impurity placed at its center. Unlike previous studies, where the overall states of the system were considered, here for the first time we have studied the fidelity for partial one- and two-site states. We have shown that the point of the quantum phase transition is still clearly marked by the sudden drop of the mixed state fidelity, even in the simplest case of the one-site partial states, as long as they are close to the impurity itself. The same holds for the two-site partial states—the fidelity exhibits a sudden drop in the vicinity of the quantum phase transition. Here, for the transition to be inferred it is necessary that one of the subsystem electrons lies at the impurity, or its neighborhood. This is expected to occur for this model since, as shown earlier, the impurity acts mainly as a local perturbation (but see also the behavior of the nonvanishing two-site "long distance" entanglement measures discussed in [35]). Such behavior of the one- and two-site fidelities, as well as the additional study of the one-site fidelity between two different sites (for a common value of the driving parameter J), enables us to identify the on-site magnetization as an optimal observable for the one-site fidelity and the order parameter for the transition considered. We have also evaluated the Uhlmann mixed state geometric phase for the cases of both one- and two-site partial states and found it to be trivial (i.e., equal to 1) for every value of the quantum phase transition driving parameter J, except when an electron site is at the impurity, thus showing that the structure of the subsystem eigenvectors has little relevance to the magnetic transition discussed. Finally, we have shown that the fidelity approach can also be useful in studying the total correlations between subsystems of the whole physical system, and how the point of quantum phase transition can be inferred from the change in the amount of total correlations [given by the quantity C_2 in (20)], which is closely related to the fidelity between the partial states discussed in this work.

We believe that this work offers yet another confirmation of the usefulness of the fidelity approach to study quantum phase transitions. It extends this approach to partial states, showing that, in some cases at least, quantum phase transitions can be signaled out by the behavior of the mixed state fidelity taken between local partial states. It is an interesting and challenging task to try to find a unified view to this approach and possibly connect it to some previously well-studied thermodynamic quantities, as was previously done for the case of global quantities (fidelity and susceptibility) in symmetry-breaking phase transitions of LGW type, see Eq. (4). The generalization to thermal phase transitions is another possible path of further research, as well as finding other phase transitions using the fidelity approach. Finally, we see our simplified study of the total correlations in a composite quantum state, based on the fidelity, as just a first step toward exploring the possibility of using the fidelity-based quantum state distinguishability for quantifying correlations, both classical as well as quantum (entanglement), a subject of research that has on its own found rich applications within the field of condensed matter and many-body physics.

ACKNOWLEDGMENTS

One of the authors (N.P.) is grateful for the support from EU FEDER and FCT Grant No. SFRH/BPD/31807/2006 and project QSec PTDC/EIA/67661/2006. The authors (P.D.S., P.N., V.R.V.) acknowledge the support of projects SFA-2-91, the ESF Science Programme INSTANS 2005-2010, PTDC/FIS/70843/2006. One of the authors (V.K.D.) is supported by the Polish Ministry of Science and Higher Education for research projects in years 2006–2009 and 2007–2010 and by STCU Grant No. 3098 in Ukraine. This work is also supported by the cooperation agreement between Poland and Portugal in the years 2007–2008.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[2] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 1999).

[3] *Strongly Correlated Systems, Coherence and Entanglement*, edited by J. M. P. Carmelo, P. D. Sacramento, J. M. B. L. dos Santos, and V. R. Vieira (World Scientific, Singapore, 2007).

[4] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
[5] P. Zanardi and N. Paunković, Phys. Rev. E **74**, 031123 (2006).
[6] P. Zanardi, M. Cozzini, and P. Giorda, J. Stat. Mech.: Theory Exp. (2007) L02002; M. Cozzini, P. Giorda, and P. Zanardi, Phys. Rev. B **75**, 014439 (2007).
[7] P. Buonsante and A. Vezzani, Phys. Rev. Lett. **98**, 110601 (2007).
[8] N. Oelkers and J. Links, Phys. Rev. B **75**, 115119 (2007).
[9] S. Chen, L. Wang, S.-J. Gu, and Y. Wang, Phys. Rev. E **76**, 061108 (2007).
[10] M.-F. Yang, Phys. Rev. B **76**, 180403(R) (2007).
[11] H.-Q. Zhou and J. P. Barjaktarevic, e-print arXiv:cond-mat/0701608; H.-Q. Zhou, J.-H. Zhao, and B. Li, e-print arXiv:0704.2940; H.-Q. Zhou, e-print arXiv:0704.2945.
[12] M. Cozzini, R. Ionicioiu, and P. Zanardi, Phys. Rev. B **76**, 104420 (2007).
[13] A. Hamma, W. Zhang, S. Haas, and D. A. Lidar, e-print arXiv:0705.0026.
[14] S.-J. Gu, H.-M. Kwok, W.-Q. Ning, and H.-Q. Lin, e-print arXiv:0706.2495.
[15] P. Zanardi, P. Giorda, and M. Cozzini, Phys. Rev. Lett. **99**, 100603 (2007).
[16] L. Campos Venuti and P. Zanardi, Phys. Rev. Lett. **99**, 095701 (2007).
[17] W.-L. You, Y.-W. Li, and S.-J. Gu, Phys. Rev. E **76**, 022101 (2007).
[18] P. Zanardi, H. T. Quan, X. Wang, and C. P. Sun, Phys. Rev. A **75**, 032109 (2007).
[19] P. Zanardi, L. Campos Venuti, and P. Giorda, Phys. Rev. A **76**, 062318 (2007).
[20] N. Paunković and V. R. Vieira, Phys. Rev. E **77**, 011129 (2008).
[21] A. C. M. Carollo and J. K. Pachos, Phys. Rev. Lett. **95**, 157203 (2005); S.-L. Zhu, *ibid.* **96**, 077206 (2006); A. Hamma, e-print arXiv:quant-ph/0602091; M. E. Reuter, M. J. Hartmann, and M. B. Plenio, Proc. R. Soc. London, Ser. A **463**, 1271 (2007).
[22] A. Uhlmann, Rep. Math. Phys. **24**, 229 (1986).
[23] C. A. Fuchs, e-print arXiv:quant-ph/9601020.
[24] A. Sakurai, Prog. Theor. Phys. **44**, 1472 (1970).
[25] K. Satori, H. Shiba, O. Sakai, and Y. Shimizu, J. Phys. Soc. Jpn. **61**, 3239 (1992).
[26] M. I. Salkola, A. V. Balatsky, and J. R. Schrieffer, Phys. Rev. B **55**, 12648 (1997).
[27] D.K. Morr and N.A. Stavropoulos, Phys. Rev. B **67**, 020502(R) (2003).
[28] D. K. Morr and J. Yoon, Phys. Rev. B **73**, 224511 (2006).
[29] A. V. Balatsky, I. Vekhter, and J.-X. Zhu, Rev. Mod. Phys. **78**, 373 (2006).
[30] A. A. Abrikosov and L. P. Gor'kov, Zh. Eksp. Teor. Fiz. **39**, 178 (1960) [Sov. Phys. JETP **12**, 1243 (1961)].
[31] R. Ramazashvili and P. Coleman, Phys. Rev. Lett. **79**, 3752 (1997).
[32] P. D. Sacramento, V. K. Dugaev, and V. R. Vieira, Phys. Rev. B **76**, 014512 (2007).
[33] P. G. de Gennes, *Superconductivity of Metals and Alloys* (Addison-Wesley, Reading, MA, 1989).
[34] J. A. M. Vermaseren, e-print arXiv:math-ph/0010025v2.
[35] P. D. Sacramento, P. Nogueira, V. R. Vieira, and V. K. Dugaev, Phys. Rev. B **76**, 184517 (2007).
[36] J. Tidström and E. Sjöqvist, Phys. Rev. A **67**, 032110 (2003).
[37] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).
[38] Note that in the case of pure states, this corresponds to the case of mutually orthogonal state vectors.
[39] The particular numerical value for δJ was chosen such that the effect of the fidelity drop becomes observable. Alternatively, a more formal approach could be taken, involving the study of a metric tensor, a local quantity induced by the global distance based on the notion of fidelity (see, for example, [15,16,19], where the Bures metric was studied).
[40] These interaction terms are easily expressed in terms of the electron spin operators $\hat{\vec{S}}$ and the Nambu operators $\hat{\vec{T}}$, discussed in detail in [20]. They commute with each other, i.e., $[\hat{T}^{\alpha},\hat{S}^{\beta}]=0$, for every $\alpha,\beta\in\{0,+,-\}$, and they have a common eigenbasis. In fact, these operators satisfy $\hat{T}^{\alpha}\hat{S}^{\beta}=0$, and act nontrivially in mutually orthogonal subspaces, with even and odd number of particle occupation numbers, respectively, annihilating the states in the other subspace.

Quantum walk on a line with two entangled particles

Y. Omar
SQIG, Instituto de Telecomunicações and CEMAPRE, ISEG, Universidade Técnica de Lisboa, P-1200-781 Lisbon, Portugal

N. Paunković
SQIG, Instituto de Telecomunicações and Instituto Superior Técnico, P-1049-001 Lisbon, Portugal

L. Sheridan
Institute for Quantum Computing, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G5

S. Bose
Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom
(Received 8 December 2005; published 3 October 2006)

We introduce the concept of a quantum walk with two particles and study it for the case of a discrete time walk on a line. A quantum walk with more than one particle may contain entanglement, thus offering a resource unavailable in the classical scenario and which can present interesting modifications on quantum walks with single particles. In this work, we show both numerically and analytically how the entanglement and the relative phase between the states describing the *coin* degree of freedom of each particle will influence the evolution of the quantum walk. In particular, the probability to find at least one particle in a certain position after N steps of the walk, as well as the average distance and the squared distance between the two particles, can be larger or smaller than the case of two unentangled particles, depending on the initial conditions we choose. This resource can then be tuned according to our needs to modify the features of a quantum walk. Experimental implementations are briefly discussed.

DOI: 10.1103/PhysRevA.74.042304 PACS number(s): 03.67.Mn, 03.65.−w, 05.30.−d, 05.40.Fb

Quantum walks, the quantum version of random walks, were first introduced in 1993 [1] and have since then been a topic of research within the context of quantum information and computation (for an introduction, see [2]). Given the superposition principle of quantum mechanics, quantum walks allow for coherent superpositions of classical random walks and, due to interference effects, can exhibit different features and offer advantages when compared to the classical case. In particular, for a quantum walk on a line, the variance after N steps is proportional to N^2, rather than N as in the classical case (see Fig. 1). Recently, several quantum algorithms with optimal efficiency were proposed based on quantum walks [3], and it was even shown that a continuous-time quantum walk on a specific graph can be used for exponential algorithmic speedup [4].

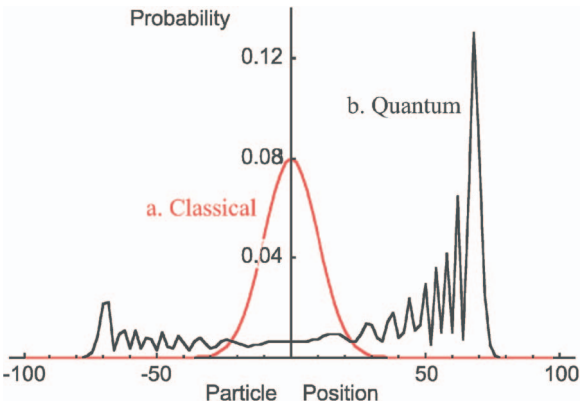


FIG. 1. (Color online) Probability distribution for a classical random walk (a) on a line after $N=100$ steps, as well as for a quantum walk (b) with initial state $|0\rangle\otimes|\uparrow\rangle$ and a Hadamard coin.

All studies on quantum walks so far have, however, been based on a single walker. In this article we study a discrete-time quantum walk on a line with two particles. Classically, random walks with K particles are equivalent to K independent single-particle random walks. In the quantum case, though, a walk with K particles may contain entanglement, thus offering a resource unavailable in the classical scenario which can introduce interesting features. Moreover, in the case of identical particles we have to take into account the effects of quantum statistics, giving an additional feature to quantum walks that can also be exploited. In this work we explicitly show that a quantum walk with two particles can indeed be tuned to behave very differently from two independent single-particle quantum walks.

Let us start by introducing the discrete-time quantum walk on a line for a single particle. The relevant degrees of freedom are the particle's position i (with $i\in\mathbb{Z}$) on the line, as well as its *coin* state. The total Hilbert space is given by $\mathcal{H}\equiv\mathcal{H}_P\otimes\mathcal{H}_C$, where $\mathcal{H}_P$ is spanned by the orthonormal vectors $\{|i\rangle\}$ representing the position of the particle and $\mathcal{H}_C$ is the two-dimensional coin space spanned by two orthonormal vectors which we denote as $|\uparrow\rangle$ and $|\downarrow\rangle$.

Each step of the quantum walk is given by two subsequent operations: the *coin operation* and the *shift-position operation*. The coin operation, given by $\hat{U}_C\in\mathrm{SU}(2)$ and acting only on $\mathcal{H}_C$, is the quantum equivalent of randomly choosing which way the particle will move (like tossing a coin in the classical case). The nonclassical character of the quantum walk is precisely here, as this operation allows for

 ©2006 The American Physical Society

superpositions of different alternatives, leading to different moves. Then, the shift-position operation $\hat{S}$ moves the particle according to the coin state, transferring this way the quantum superposition to the total state in $\mathcal{H}$. The evolution of the system at each step of the walk can then be described by the total unitary operator

$$\hat{U} \equiv \hat{S}(\hat{I}_P \otimes \hat{U}_C), \tag{1}$$

where $\hat{I}_P$ is the identity operator on $\mathcal{H}_P$. Note that if a measurement is performed after each step, we will revert to the classical random walk.

In this article we choose to study a quantum walk with a *Hadamard coin*—i.e., where $\hat{U}_C$ is the Hadarmard operator $\hat{H}$:

$$\hat{H} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{2}$$

Note that this represents a *balanced coin*; i.e., there is a 50-50 chance for each alternative. The shift-position operator is given by

$$\hat{S} = \left(\sum_i |i+1\rangle\langle i|\right) \otimes |\uparrow\rangle\langle\uparrow| + \left(\sum_i |i-1\rangle\langle i|\right) \otimes |\downarrow\rangle\langle\downarrow|. \tag{3}$$

Therefore, if the initial state of our particle is, for instance, $|0\rangle \otimes |\uparrow\rangle$, the first step of the quantum walk will be as follows:

$$|0\rangle \otimes |\uparrow\rangle \xrightarrow{\hat{H}} |0\rangle \otimes \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \xrightarrow{\hat{S}} \frac{1}{\sqrt{2}}(|1\rangle \otimes |\uparrow\rangle + |-1\rangle \otimes |\downarrow\rangle). \tag{4}$$

We see that there is a probability of 1/2 to find the particle in position 1, as well as to find it in position 2, just like in the classical case. Yet, if we let this quantum walk evolve for three or more steps before we perform a position measurement, we will find a very different probability distribution for the position of the particle when compared to the classical random walk, as can be seen in Fig. 1 for $N=100$ steps.

Let us consider the previous quantum walk, but now with two noninteracting particles on a line (not necessarily the same). If the particles are distinguishable and in a pure separable state, the position measurement of one particle will not change the probability distribution of the other; they are completely uncorrelated. On the other hand, if the particles are entangled, a new resource with no classical equivalent will be at our disposal.

This is a specific instance of the well-known interference phenomenon in quantum mechanics, applied to the case of a composite system. In particular, let us have a composite system consisting of two particles, labeled 1 and 2, prepared in a pure separable state,

$$|\psi^S\rangle_{12} = |\alpha\rangle_1 |\beta\rangle_2, \tag{5}$$

and a general separable two-particle observable

$$\hat{A}_{12} = \hat{A}_1 \otimes \hat{A}_2. \tag{6}$$

Then, the expectation value of the observable $\hat{A}_{12}$ for a system in a state $|\psi^S\rangle_{12}$ is the simple product of the two subsystem expectation values:

$$\langle\psi^S|\hat{A}_{12}|\psi^S\rangle_{12} = \langle\alpha|\hat{A}_1|\alpha\rangle_1 \langle\beta|\hat{A}_2|\beta\rangle_2. \tag{7}$$

On the other hand, if the composite system is in a general entangled state

$$|\chi^E\rangle_{12} = a|\psi^S\rangle_{12} + b|\varphi^S\rangle_{12}, \tag{8}$$

where $|a|^2+|b|^2=1$ and $|\varphi^S\rangle_{12}=|\beta\rangle_1|\alpha\rangle_2$ (with $|\alpha\rangle \neq e^{i\theta}|\beta\rangle$), then the expectation value of the observable $\hat{A}_{12}$ is given by

$$\begin{aligned}\langle\chi^E|\hat{A}_{12}|\chi^E\rangle_{12} = {} & |a|^2\langle\alpha|\hat{A}_1|\alpha\rangle_1\langle\beta|\hat{A}_2|\beta\rangle_2 \\ & + |b|^2\langle\beta|\hat{A}_1|\beta\rangle_1\langle\alpha|\hat{A}_2|\alpha\rangle_2 \\ & + a^*b\langle\alpha|\hat{A}_1|\beta\rangle_1\langle\beta|\hat{A}_2|\alpha\rangle_2 \\ & + ab^*\langle\beta|\hat{A}_1|\alpha\rangle_1\langle\alpha|\hat{A}_2|\beta\rangle_2.\end{aligned} \tag{9}$$

Note that the first two terms on the right-hand side of the above equation represent the weighted sum of the expectation values for the states $|\psi^S\rangle_{12}$ and $|\varphi^S\rangle_{12}$, with the weights $|a|^2$ and $|b|^2$, respectively. In other words, the first two terms represent the expectation value of the observable $\hat{A}_{12}$ for a mixed separable state:

$$\hat{\rho}^S_{12} = |a|^2|\psi^S\rangle\langle\psi^S|_{12} + |b|^2|\varphi^S\rangle\langle\varphi^S|_{12}. \tag{10}$$

The last two terms in Eq. (9) are purely quantum and represent the interference between the two states $|\psi^S\rangle_{12}$ and $|\varphi^S\rangle_{12}$, superposed in the state $|\chi\rangle_{12}$. Note that to obtain this purely quantum phenomenon, we need only an entangled state, even if the observable $\hat{A}_{12}=\hat{A}_1\otimes\hat{A}_2$ is separable.

Returning to the case of two walkers on a line, the joint Hilbert space of our composite system is given by

$$\mathcal{H}_{12} \equiv \mathcal{H}_1 \otimes \mathcal{H}_2 \equiv (\mathcal{H}_{P,1} \otimes \mathcal{H}_{C,1}) \otimes (\mathcal{H}_{P,2} \otimes \mathcal{H}_{C,2}), \tag{11}$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ represent the Hilbert spaces of particles 1 and 2, respectively. Since the relevant degrees of freedom in our problem are the same for both particles, we have that both $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic to $\mathcal{H}$ defined earlier for the one-particle case. Note also that in the case of identical particles we have to restrict $\mathcal{H}_{12}$ to its symmetrical and antisymmetrical subspaces for bosons and fermions, respectively.

Let us first consider the particular case where both particles start the quantum walk in the same position, 0, but with different coin states $|\downarrow\rangle$ and $|\uparrow\rangle$. In the case where the particles are in a pure separable state, our system's initial state will be given by

$$|\psi^S_0\rangle_{12} = |0,\downarrow\rangle_1|0,\uparrow\rangle_2. \tag{12}$$

We can also consider an initial pure state entangled in the coin degrees of freedom. In particular, we will consider the following two maximally entangled states:

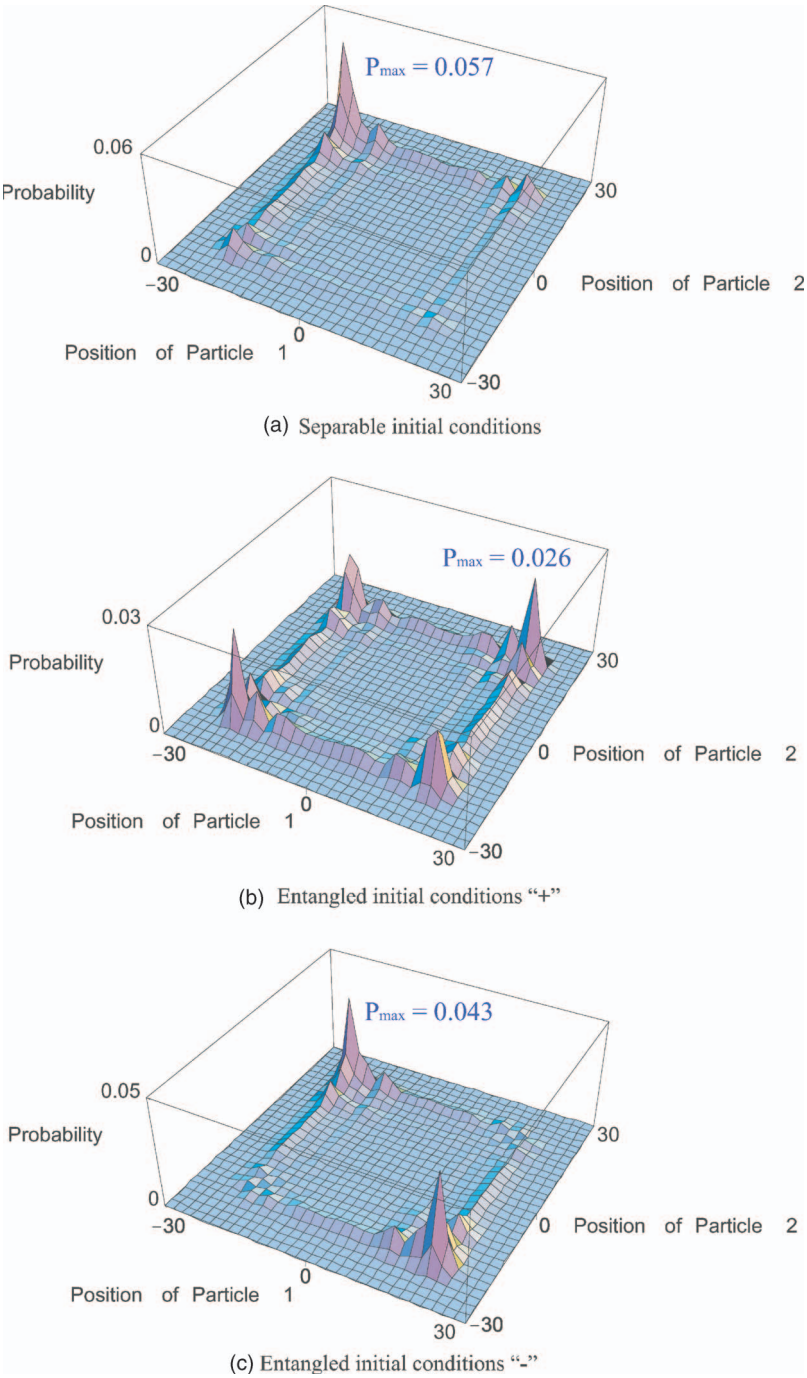


FIG. 2. (Color online) Two-particle probability distributions after $N=30$ steps for different initial conditions: (a) separable state $|\psi_0^S\rangle_{12}$, (b) $|\psi_0^+\rangle_{12}$ state, and (c) $|\psi_0^-\rangle_{12}$ state. Note the different vertical ranges.

$$|\psi_0^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|0,\downarrow\rangle_1|0,\uparrow\rangle_2 \pm |0,\uparrow\rangle_1|0,\downarrow\rangle_2), \tag{13}$$

differing only by a relative phase. Note that if we were considering identical particles on the same point, our system would have to be described by these states for bosons and fermions, respectively.

Each step of this two-particle quantum walk will be given by

$$\hat{U}_{12} = \hat{U} \otimes \hat{U}, \tag{14}$$

where $\hat{U}$ is given by Eq. (1) and is the same for both particles. After N steps, the state of the system will be, in the case of the initial conditions (12)

$$|\psi_N^S\rangle_{12} = \hat{U}_{12}^N|\psi_0^S\rangle_{12} = \hat{U}^N|0,\downarrow\rangle_1\hat{U}^N|0,\uparrow\rangle_2. \tag{15}$$

Figure 2(a) shows the joint probability distribution $P_{12}^S(i,j;N)$ for finding particle 1 in position i and particle 2 in position j for $N=30$ steps. Note that, since the particles are

uncorrelated, $P_{12}^{S}(i,j;N)$ is simply the product of the two independent one-particle distributions:

$$P_{12}^{S}(i,j;N) = P_{1}^{S}(i;N) \times P_{2}^{S}(j;N), \tag{16}$$

where $P_{1}^{S}(i;N)$ is the probability distribution for finding particle 1 in position i after N steps and similarly for $P_{2}^{S}(j;N)$ and particle 2. This can also be observed in Fig. 2(a), which is clearly the product of two independent single-particle distributions, one biased to the left for particle 1 and the other to the right for particle 2, accordingly with the initial conditions given by Eq. (12).

In the case of entangled particles, the state of the system after N steps will be

$$|\psi_{N}^{\pm}\rangle_{12} = \hat{U}_{12}^{N}|\psi_{0}^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(\hat{U}^{N}|0,\downarrow\rangle_{1}\hat{U}^{N}|0,\uparrow\rangle_{2} \pm \hat{U}^{N}|0,\uparrow\rangle_{1}\hat{U}^{N}|0,\downarrow\rangle_{2}). \tag{17}$$

The probability distribution for finding particle 1 in position i and particle 2 in position j in the "+" case, $P_{12}^{+}(i,j;N)$, is represented in Fig. 2(b), for $N=30$. Similarly, Fig. 2(c) shows the distribution $P_{12}^{-}(i,j;N)$ of the "−" case, again for $N=30$. The effects of the entanglement are striking when comparing the three distributions in Fig. 2. In particular, we see that these effects significantly increase the probability of the particles reaching certain configurations on the line, which otherwise would be very unlikely to be occupied. In all cases, the maxima of the distributions occur around positions $\pm N/\sqrt{2} \approx \pm 20$. In the "+" case it is most likely to find both particles together, whereas for "−" the former situation is impossible and the particles will tend to finish as distant as possible from one another.

Let us now consider the individual particles in the entangled system. For instance, the state of the first particle can be described by the reduced density operator $\hat{\rho}_{1}(N) \equiv Tr_{2}(|\psi_{N}^{\pm}\rangle_{12\ 12}\langle\psi_{N}^{\pm}|)$, which consists of an equal mixture of the one-particle states $\hat{U}^{N}|0,\downarrow\rangle$ and $\hat{U}^{N}|0,\uparrow\rangle$. Thus, the probability to find particle 1 in position i after N steps is given by the following marginal probability distribution:

$$P_{1}^{\pm}(i;N) = \frac{1}{2}[P_{\downarrow}(i;N) + P_{\uparrow}(i;N)], \tag{18}$$

where $P_{\downarrow}(i;N)$ is the probability distribution for finding the particle in state $\hat{U}^{N}|0,\downarrow\rangle$ in position i after N steps and similarly for $P_{\uparrow}(i;N)$ and a particle in state $\hat{U}^{N}|0,\uparrow\rangle$. The results for particle 2 are analogous: in particular, $P_{2}^{\pm}(i;N)$ is also given by Eq. (18). Note that in the case of initial conditions (12) we have the marginal probabilities $P_{1}^{S}(i;N)=P_{\downarrow}(i;N)$ and $P_{2}^{S}(i;N)=P_{\uparrow}(i;N)$. But now, contrary to the separable case, the joint probability $P_{12}^{\pm}(i,j;N)$ is no longer the simple product of the two one-particle probabilities. In fact, it is not even a weighted sum of productlike terms as it contains the interference part analogous to the one from Eq. (9), giving information about the nontrivial *quantum* correlations between the outcomes of the position measurement performed on each particle. From Eq. (17), it follows that the probability $P_{12}^{\pm}(i,j;N)$ is given by

$$P_{12}^{\pm}(i,j;N) = \frac{1}{2}\{P_{\downarrow\uparrow}(i,j;N) + P_{\uparrow\downarrow}(i,j;N) \pm [\mathcal{I}_{\downarrow\uparrow}(i;N)\mathcal{I}_{\uparrow\downarrow}(j;N) + \mathcal{I}_{\uparrow\downarrow}(i;N)\mathcal{I}_{\downarrow\uparrow}(j;N)]\}, \tag{19}$$

where $P_{\downarrow\uparrow}(i,j;N)=P_{\downarrow}(i;N)P_{\uparrow}(j;N)$ and $\mathcal{I}_{\downarrow\uparrow}(i;N) = \langle 0,\downarrow|(\hat{U}^{\dagger})^{N}|i\rangle\langle i|\hat{U}^{N}|0,\uparrow\rangle$ [and analogously for $P_{\uparrow\downarrow}(i,j;N)$ and $\mathcal{I}_{\downarrow\uparrow}(i;N)$]. Therefore, to investigate more quantitatively the difference between quantum walks with two distinguishable particles in a pure separable state and two particles entangled in their coin degree of freedom, we must look at joint (two-particle) rather than individual properties.

TABLE I. Average product $\langle\hat{x}_1\hat{x}_2\rangle^{S,\pm}$ after N steps, for the initial conditions $|\psi_0^-\rangle_{12}$, $|\psi_0^S\rangle_{12}$, and $|\psi_0^+\rangle_{12}$.

	Expectation value $\langle\hat{x}_1\hat{x}_2\rangle^{S,\pm}$ after N steps					
N	10	20	30	40	60	100
$\|\psi_0^-\rangle_{12}$	−16.8	−69.8	−153.5	−276.2	−619.7	−1718.3
$\|\psi_0^S\rangle_{12}$	−6.0	−31.3	−69.9	−130.5	−298.3	−839.6
$\|\psi_0^+\rangle_{12}$	4.8	7.3	13.7	15.1	23.1	39.1

First, let us consider a simple product $\hat{x}_1\hat{x}_2$ between the subsystem position operators $\hat{x}_1$ and $\hat{x}_2$ for particles 1 and 2, respectively (note that for simplicity, we have omitted to write the dependence on the number of steps, N). The expectation values $\langle\hat{x}_1\hat{x}_2\rangle^{S,\pm}$ for the three initial states $|\psi_0^S\rangle_{12}$ and $|\psi_0^{\pm}\rangle_{12}$ are presented in Table I for different N. We observe that the three cases are "equidistant" in the sense that, for each N, the following relation holds:

$$\langle\hat{x}_1\hat{x}_2\rangle^{-} - \langle\hat{x}_1\hat{x}_2\rangle^{S} = \langle\hat{x}_1\hat{x}_2\rangle^{S} - \langle\hat{x}_1\hat{x}_2\rangle^{+}. \tag{20}$$

In order to study the above quantity analytically, let us write the definition of the expectation value $\langle\hat{x}_1\hat{x}_2\rangle \equiv \Sigma_{i,j=-N}^{N} ij P_{12}(i,j;N)$. Using Eqs. (16) and (19) for the joint probability distributions $P_{12}^{S}(i,j;N)$ and $P_{12}^{\pm}(i,j;N)$, respectively, we get the following expression:

$$\langle\hat{x}_1\hat{x}_2\rangle^{\pm} = \langle\hat{x}_1\hat{x}_2\rangle^{S} \pm \frac{1}{2}\sum_{i,j=-N}^{N} ij[\mathcal{I}_{\downarrow\uparrow}(i;N)\mathcal{I}_{\uparrow\downarrow}(j;N) + \mathcal{I}_{\uparrow\downarrow}(i;N)\mathcal{I}_{\downarrow\uparrow}(j;N)]. \tag{21}$$

We immediately see that the relation (20) is indeed satisfied for every N. If by $\langle\hat{x}_{\downarrow}\rangle$ we denote the average position for the one-particle probability distribution $P_{\downarrow}(i;N)$ and analogously for $\langle\hat{x}_{\uparrow}\rangle$, we also have that $\langle\hat{x}_1\hat{x}_2\rangle^{S}=\langle\hat{x}_{\downarrow}\rangle\langle\hat{x}_{\uparrow}\rangle=-\langle\hat{x}_{\downarrow}\rangle^2$ as $P_{\downarrow}(i;N)=P_{\uparrow}(-i;N)$. Using the obvious relation $\mathcal{I}_{\downarrow\uparrow}(i;N) = \mathcal{I}_{\uparrow\downarrow}^{*}(i;N)$, we can further simplify expression (21) by writing

$$\langle\hat{x}_1\hat{x}_2\rangle^{\pm} = -\langle\hat{x}_{\downarrow}\rangle^2 \pm \|I_{\downarrow\uparrow}\|^2, \tag{22}$$

with $I_{\downarrow\uparrow}=\Sigma_{i=-N}^{N} i\mathcal{I}_{\downarrow\uparrow}(i;N)$.

To analyze the asymptotic behavior of $\langle\hat{x}_1\hat{x}_2\rangle$ for large N,

we use the technique developed in [6] based on Fourier analysis (for an alternative treatment, see [7]). There, it was shown that for every initial coin state, the expectation value of the position operator in a one-particle quantum walk on a line has a linear drift $\langle \hat{x} \rangle \propto N$, for N large enough. Following [6], we can express $I_{\downarrow\uparrow}$ as

$$I_{\downarrow\uparrow} = -\frac{1}{2\pi}\sum_{j=1}^{N}\int_{-\pi}^{\pi} dk \langle\downarrow|(\hat{H}_k^{\dagger})^j \hat{Z}(\hat{H}_k)^j|\uparrow\rangle, \tag{23}$$

with $\hat{Z} \equiv \hat{I}_C - 2|\downarrow\rangle\langle\downarrow|$ and

$$\hat{H}_k = \frac{1}{\sqrt{2}}\begin{bmatrix} e^{-ik} & e^{-ik} \\ e^{ik} & -e^{ik} \end{bmatrix}.$$

Let $|\phi_{kl}\rangle$ be eigenvectors for the operators $\hat{H}_k$, with eigenvalues $e^{\theta_{kl}}$, and let $|\uparrow\rangle = \Sigma_{l=1}^{2} u_{kl}|\phi_{kl}\rangle$ and $|\downarrow\rangle = \Sigma_{l=1}^{2} d_{kl}|\phi_{kl}\rangle$. Then, Eq. (23) can be rewritten in the following way ($\hat{P}_{\downarrow} \equiv |\downarrow\rangle\langle\downarrow|$):

$$I_{\downarrow\uparrow} = -N + \frac{1}{\pi}\int_{-\pi}^{\pi} dk \sum_{l,l'=1}^{2} u_{kl}^{*} d_{kl'} \langle\phi_{kl}|\hat{P}_{\downarrow}|\phi_{kl'}\rangle \sum_{j=1}^{N} e^{i(\theta_{kl'}-\theta_{kl})j}. \tag{24}$$

Since for $l \neq l'$ the terms in the above equation are oscillatory, for large N they will average to zero, thus yielding the linear asymptotic behavior of $I_{\downarrow\uparrow}$ with respect to the number of steps:

$$I_{\downarrow\uparrow} = E_1 N, \tag{25}$$

where $E_1 = -1 + (1/\pi)\int_{-\pi}^{\pi} dk \Sigma_l u_{kl}^{*} d_{kl} \langle\phi_{kl}|\hat{P}_{\downarrow}|\phi_{kl}\rangle$. Combining the above result with Eq. (22) and the asymptotic behavior $\langle \hat{x} \rangle \propto N$, we conclude that in general $\langle \hat{x}_1 \hat{x}_2 \rangle$, as well as the "distances" given by Eq. (20), should scale quadratically with the number of steps N. Looking at the numerical results presented in Table I, we see that for $|\psi_0^{-}\rangle_{12}$ and $|\psi_0^{S}\rangle_{12}$ initial states we indeed have that $\langle \hat{x}_1 \hat{x}_2 \rangle^{S,-} \propto N^2$, while for the $|\psi_0^{+}\rangle_{12}$ initial state we have that $\langle \hat{x}_1 \hat{x}_2 \rangle^{+} \propto N$. In other words, the quadratic parts of $\langle \hat{x}_{\downarrow} \rangle^2$ and $\|I_{\downarrow\uparrow}\|^2$ cancel out each other for the case of $|\psi_0^{+}\rangle_{12}$ initial state. We performed numerical simulations up to $N=8000$ steps, and they confirmed the behavior presented in Table I.

Note that these effects require the presence of entanglement in the initial conditions. For instance, $P_{12}^{-}(i,j;N)$ can never be obtained from *separable* initial states of the coins. Suppose that, given a general separable (possibly mixed) initial state of the coins, one has to achieve in N steps as negative a value of $\langle \hat{x}_1 \hat{x}_2 \rangle$ as possible by starting the walkers in the position state $|0\rangle_1|0\rangle_2$. For such states, $\langle \hat{x}_1 \hat{x}_2 \rangle$ is simply a weighted average of $\langle \hat{x}_1 \rangle \langle \hat{x}_2 \rangle$ over several pure product states. Now, it is well known that for single-particle walks with the Hadamard coin, $|0,\uparrow\rangle_i$ and $|0,\downarrow\rangle_i$ give the most positive and the most negative values of $\langle \hat{x}_i \rangle$, respectively [2] (they give the most asymmetric distributions for the walkers' position). Thus, for two-particle walks, a set of separable states for most negative $\langle \hat{x}_1 \hat{x}_2 \rangle$ are mixed states of the form $\hat{\rho}^A = \mathcal{P}|\psi_0^S\rangle\langle\psi_0^S| + (1-\mathcal{P})|\psi_0^{S'}\rangle\langle\psi_0^{S'}|$, where $|\psi_0^{S'}\rangle_{12} = |0,\uparrow\rangle_1|0,\downarrow\rangle_2$ and $\mathcal{P}$ is any probability. Mixing *any* state $|0,\nearrow\rangle_1|0,\searrow\rangle_2$ with $\hat{\rho}^A$, where $|\nearrow\rangle$ and $|\searrow\rangle$ are arbitrary states of the coins (not necessarily orthogonal to each other), even with an arbitrarily small probability, will make $\langle \hat{x}_1 \hat{x}_2 \rangle$ more positive or keep it unchanged at best. However, from Eqs. (21) and (22), we see that $\langle \hat{x}_1 \hat{x}_2 \rangle^{-} < \langle \hat{x}_1 \hat{x}_2 \rangle^{S}$ for every N. Thus, the expectation value $\langle \hat{x}_1 \hat{x}_2 \rangle$ of $P_{12}^{-}(i,j;N)$ is always more negative than that obtained with *any* separable state. Therefore, $P_{12}^{-}(i,j;N)$ cannot be reproduced with a walk starting in $|0\rangle_1|0\rangle_2$ that has a separable initial state of the coins.

For a general pure entangled coin state we can write the overall initial state as follows:

$$|\psi_0\rangle_{12} = |00\rangle_{12} \otimes (\alpha|\uparrow\uparrow\rangle_{12} + \beta|\uparrow\downarrow\rangle_{12} + \gamma|\downarrow\uparrow\rangle_{12} + \delta|\downarrow\downarrow\rangle_{12}), \tag{26}$$

with $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. From the above discussion, it follows that the expectation value of operator $\hat{x}_1 \hat{x}_2$ can be written in the following form:

$$\begin{aligned} \langle \hat{x}_1 \hat{x}_2 \rangle = {} & (|\alpha|^2 + |\delta|^2 - |\beta|^2 - |\gamma|^2)\langle \hat{x}_{\downarrow} \rangle^2 \\ & + 2\langle \hat{x}_{\downarrow} \rangle \mathrm{Re}[(-\alpha\beta^{*} - \alpha\gamma^{*} + \gamma\delta^{*} + \beta\delta^{*})I_{\downarrow\uparrow}] \\ & + 2\,\mathrm{Re}[(\alpha\delta^{*} + \beta\gamma^{*})I_{\downarrow\uparrow}^2], \end{aligned} \tag{27}$$

where $\mathrm{Re}[z]$ is real part of a complex number z. The initial entangled states $|\psi_0^{\pm}\rangle_{12}$ discussed above correspond to the choice of coefficients $\alpha = \delta = 0$ and $\gamma = \pm\beta = 1/\sqrt{2}$. We see that by taking as initial coin states the other two Bell states (corresponding to the choice of $\beta = \gamma = 0$ and $\delta = \pm\alpha = 1/\sqrt{2}$)

$$|\phi_0^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|0,\downarrow\rangle_1|0,\downarrow\rangle_2 \pm |0,\uparrow\rangle_1|0,\uparrow\rangle_2), \tag{28}$$

we obtain analogous results: $|\phi_0^{\pm}\rangle_{12}$ are "equidistant" from the separable initial states $|0,\downarrow\rangle_1|0,\downarrow\rangle_2$ and $|0,\uparrow\rangle_1|0,\uparrow\rangle_2$ [in the sense of Eq. (20)], with $|\phi_0^{+}\rangle_{12}$ achieving the maximal positive value for $\langle \hat{x}_1 \hat{x}_2 \rangle$, which has the same absolute value as the one obtained for the $|\psi_0^{-}\rangle_{12}$ initial state.

Finally, we note that the above results for $\langle \hat{x}_1 \hat{x}_2 \rangle$ [Eq. (22) and Table I], together with the results for $\langle \hat{x} \rangle$ of a single-particle quantum walk on a line (see [6]), offer us an explicit form for the correlation function $C_{12} \equiv \langle \hat{x}_1 \hat{x}_2 \rangle - \langle \hat{x}_1 \rangle \langle \hat{x}_2 \rangle$.

The second quantity we choose to study is the squared distance between the two particles, defined as

$$\langle \hat{\Delta}_{12}^2 \rangle \equiv \langle (\hat{x}_1 - \hat{x}_2)^2 \rangle. \tag{29}$$

Again, it represents a joint property. The expectation values $\langle \hat{\Delta}_{12}^2 \rangle^{S,\pm}$ for the three initial states $|\psi_0^{S}\rangle_{12}$ and $|\psi_0^{\pm}\rangle_{12}$ are presented in Table II for different N. We see that in the "$-$" case the particles tend, on average, to end the quantum walk more distant from each other, whereas in the "$+$" case they tend to stay closer and somewhere in between in the separable case. In fact, for fixed N, we have a relation analogous to Eq. (20):

$$\langle \hat{\Delta}_{12}^2 \rangle^{-} - \langle \hat{\Delta}_{12}^2 \rangle^{S} = \langle \hat{\Delta}_{12}^2 \rangle^{S} - \langle \hat{\Delta}_{12}^2 \rangle^{+}. \tag{30}$$

This can be shown to be a general property if we express $\langle \hat{\Delta}_{12}^2 \rangle$ in the following way:

TABLE II. Average squared distance $\langle\hat{\Delta}_{12}^2\rangle^{S,\pm}$ after N steps, for the initial conditions $|\psi_0^-\rangle_{12}$, $|\psi_0^S\rangle_{12}$, and $|\psi_0^+\rangle_{12}$.

	Expectation value $\langle\hat{\Delta}_{12}^2\rangle^{S,\pm}$ after N steps					
N	10	20	30	40	60	100
$\|\psi_0^-\rangle_{12}$	93.6	374.8	835.4	1490.5	3349.1	9295.4
$\|\psi_0^S\rangle_{12}$	71.9	297.7	668.2	1199.2	2706.4	7538.0
$\|\psi_0^+\rangle_{12}$	50.3	220.6	501.1	907.9	2063.6	5780.6

$$\langle\hat{\Delta}_{12}^2\rangle = \langle\hat{x}_1^2\rangle + \langle\hat{x}_2^2\rangle - 2\langle\hat{x}_1\hat{x}_2\rangle. \tag{31}$$

Using the previous discussion for $\langle\hat{x}_1\hat{x}_2\rangle$, we see that the relation (30) is just a consequence of the result (22), as in all three cases we have that $\langle\hat{x}_1^2\rangle=\langle\hat{x}_2^2\rangle=\langle\hat{x}_\downarrow^2\rangle$, with $\langle\hat{x}_\downarrow^2\rangle$ being the mean-squared position for the probability distribution $P_\downarrow(i;N)$. In the case of a general initial coin state (26), the analytical result for $\langle\hat{\Delta}_{12}^2\rangle$ is obtained using Eq. (27) together with

$$\langle\hat{x}_1^2\rangle = \langle\hat{x}_\downarrow^2\rangle + 2\,\mathrm{Re}[(\alpha\beta^* + \gamma\delta^*)K_{\downarrow\uparrow}], \tag{32}$$

where $K_{\downarrow\uparrow}=\Sigma_{i=-N}^{N} i^2\langle 0,\downarrow|(\hat{U}^\dagger)^N|i\rangle\langle i|\hat{U}^N|0,\uparrow\rangle$ (and the analogous expression for $\langle\hat{x}_2^2\rangle$). Then, further analysis of the quantity $K_{\downarrow\uparrow}$ can be carried out in a similar fashion to the one done for $I_{\downarrow\uparrow}$ (see also [6]). The main conclusion of this analysis is that, asymptotically, $\langle\hat{\Delta}_{12}^2\rangle$ scales quadratically with the number of steps, N, which is confirmed by numerics as well; see Table II (again, we performed numerical simulations up to $N=8000$ steps, confirming the results presented in Table II).

If instead of the squared distance we choose the linear distance itself, given by $\langle\hat{\Delta}_{12}\rangle=\langle|\hat{x}_1-\hat{x}_2|\rangle$, we obtain the results presented in Table III. Again, we observe behavior analogous to that given by Eqs. (20) and (30). Furthermore, we see that the distance scales linearly with N.

Finally, let us now calculate, for the different initial conditions, the probability of finding at least one particle in position i after N steps: $\mathcal{P}^{S,\pm}(i;N)$. This is clearly a joint property as it depends on both one-particle outcomes:

$$\begin{aligned}\mathcal{P}^{S,\pm}(i;N) &= \sum_{j=-N}^{N}[P_{12}^{S,\pm}(i,j;N) + P_{12}^{S,\pm}(j,i;N)] - P_{12}^{S,\pm}(i,i;N)\\ &= [P_1^{S,\pm}(i;N) + P_2^{S,\pm}(i;N)] - P_{12}^{S,\pm}(i,i;N)\\ &= [P_\downarrow(i;N) + P_\uparrow(i;N)] - P_{12}^{S,\pm}(i,i;N).\end{aligned} \tag{33}$$

TABLE III. Average distance $\langle\hat{\Delta}_{12}\rangle^{S,\pm}$ after N steps, for the initial conditions $|\psi_0^-\rangle_{12}$, $|\psi_0^S\rangle_{12}$, and $|\psi_0^+\rangle_{12}$.

	Expectation value $\langle\hat{\Delta}_{12}\rangle^{S,\pm}$ after N steps					
N	10	20	30	40	60	100
$\|\psi_0^-\rangle_{12}$	8.8	17.5	26.0	34.9	52.2	87.0
$\|\psi_0^S\rangle_{12}$	7.1	14.7	21.9	29.5	44.3	73.9
$\|\psi_0^+\rangle_{12}$	5.5	11.9	17.8	24.1	36.3	60.8

From Eqs. (16) and (19), it follows that $P_{12}^{\pm}(i,i;N) = P_{12}^{S}(i,i;N) \pm \|\mathcal{I}_{\downarrow\uparrow}(i,i;N)\|^2$, so that

$$\mathcal{P}^-(i;N) > \mathcal{P}^S(i;N) > \mathcal{P}^+(i;N). \tag{34}$$

We see that, by introducing entanglement in the initial conditions of our two-particle quantum walk, the probability of finding at least one particle in a particular position on the line can actually be better or worse than in the case where the two particles are independent. Note that in this case this does not depend on the particular amount of entanglement introduced, as both states in Eq. (13) are maximally entangled, but rather on their symmetry-relative phase.

This work allows for generalizations in a number of ways. First, one could consider periodic or other boundary conditions on the line or more general graphs [5]. Note that the positions on the line of our two particles could also be interpreted as the position of a single particle doing the quantum walk on a regular two-dimensional lattice. More general coins could also be considered [8–10], including entangling and nonbalanced coins, as well as different initial states. One could also augment the number of particles and study these quantum walks in continuous time or in their asymptotic limit. Furthermore, also very interesting and promising is to investigate the use of multiparticle quantum walks in solving mathematical or practical problems that could be encoded as a quantum walk, such as the estimation of the volume of a convex body [11] or the connectivity in a P2P network [12].

Finally, we present some brief comments about implementations of our two-particle quantum walk on a line. The methods recently proposed for the single-particle case using cavity QED [13], optical lattices [14], or ion traps [15] could be adapted to our two-particle case. For instance, in the latter we could encode the coin states in the electronic levels of two ions and the position in their *center-of-mass* or *stretch* motional modes: the coin flipping could then be obtained with a $\pi/2$ Raman pulse and the shift with a conditional optical dipole force [16]. Another possibility is to send two photons through a tree of balanced beam splitters which implement both the coin flipping and the conditional shift, again generalizing a scheme proposed for a single particle [8,17]. Note that this could be implemented with other particles as well—e.g., electrons—using a device equivalent to a beam splitter [18]. Also very interesting, regardless of any particular method or technology, is the possibility of using two indistinguishable particles on the same line to implement our quantum walk. Say we encode the coin degrees of freedom in the polarization of two photons or in the spin of two electrons: if the two particles start in the same position 0, then they will be forced to be in the states given by Eq. (13) ("+" for bosons and "−" for fermions). Although the particles will initially be only entangled in the mathematical sense (as they cannot be addressed to extract quantum correlations), this is a perfectly valid way of preparing the $|\psi_0^+\rangle_{12}$ and $|\psi_0^-\rangle_{12}$ initial states for our quantum walk, saving us the trouble of generating entangled pairs. Thus, the indistinguishability of identical particles appears as a resource, much in the same manner as it can play an useful role in quantum-information processing [19]. Here, in particular, the indistinguishability of identical particles offers a way to simplify the

preparation of the initial states for a two-particle quantum walk on a line.

The fermionic and bosonic nature of identical particles in quantum mechanics can also be used to *explain* the behavior we have obtained above for our quantum walk with two entangled particles on the same line. Although we consider our particles to be distinguishable, the specific entangled states $|\psi_0^{\pm}\rangle_{12}$ they start in are either symmetric or antisymmetric with respect to the exchange of the indices of the particles. Both the $\hat{H}$ and the $\hat{S}$ operations can be applied through global fields, as they act on both the coins. In other words, there is no need to distinguish the coins through their labels 1 and 2 to accomplish the quantum walk we describe, and the particles could as well have been indistinguishable. One would thus get precisely the *same* results as the "−" and "+" cases if one started with two fermions at $|0\rangle$ and two bosons at $|0\rangle$, respectively. The two cases exhibit typically fermionic and bosonic behaviors, with the two particles, respectively, avoiding each other or tending to bunch together, as imposed by the symmetrization postulate of quantum mechanics [20], during their quantum walk on the line.

In this article we introduced the concept of a quantum walk with two particles and studied it both numerically and analytically for the case of a discrete-time walk on a line. Having more than one particle, we could now add a new feature to the walk: entanglement between the particles. For a general initial state of two particles starting both in position 0, we found analytical expressions for $\langle \hat{x}_1 \hat{x}_2 \rangle$ and the average squared distance and studied their asymptotic behavior. In particular, we considered initial states maximally entangled in the coin degrees of freedom and with opposite symmetries and compared them to a case where the two particles were initially unentangled. We found that the entanglement in the coin states introduced spatial correlations between the particles and that their average distance is larger in the "−" case than in the unentangled case and is smaller in the "+" case. We justified that no separable state (pure or mixed) can ever generate the probability distribution obtained for the "−" case, which shows that quantum walks with two entangled particles can offer features that are unattainable in the separable case. If it was given *a priori* that the two walkers need to hit (meaning "reach with a significant probability") two sites along the line which are anticorrelated (i.e., $x_1 = -x_2$) with a certain probability (but not with unit probability, so that one still needs to search for both the sites), then a two-particle walk starting in the "−" state could do better than any separable state. The entanglement in the initial conditions thus appears as a resource that we can tune according to our needs to enhance a given application based on a quantum walk. The possibility of also playing a role in the development of quantum algorithms seems worthwhile exploring.

The authors would like to thank K. Banaszek, E. Kashefi, and T. Schaetz for precious discussions, as well as P. Mateus, A. Andrić, and R. Chadha for their help in numerical simulations. Y.O. and N.P. gratefully acknowledge support from Fundação para a Ciência e a Tecnologia (Portugal)—namely, through programs POCTI/POCI and CLC project POCI/MAT/55796/2004 QuantLog, partially funded by FEDER and the 3rd Community Support Framework of the European Social Fund. The research of Y.O. and N.P. is also developed under the scope of the cooperation agreement 24P-22335/04 between CLE-UNICAMP and CLC-UTL. Y.O. wishes to thank the Clarendon Laboratory for its hospitality. N.P. thanks the financial support of Elsag S.p.A. and of the European Commission, Contract No. IST-2001-39215 TOPQIP. L.S. gratefully acknowledges the support of the Nuffield Foundation. Finally, Y.O. and S.B. would like to acknowledge support from CRUP/BC Collaboration Grant No. B-22/05 and thank the Institute for Quantum Information at Caltech, where this work was started, for its hospitality.

[1] Y. Aharonov, L. Davidovich, and N. Zagury, Phys. Rev. A **48**, 1687 (1993).
[2] J. Kempe, Contemp. Phys. **44**, 307 (2003).
[3] N. Shenvi, J. Kempe, and K. Bingitta Whaley, Phys. Rev. A **67**, 052307 (2003); A. Ambainis, e-print quant-ph/0311001.
[4] A. M. Childs *et al.* (unpublished).
[5] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani (unpublished); T. D. Mackay, S. D. Bartlett, L. T. Stephenson, and B. C. Sanders, J. Phys. A **35**, 2745 (2002); B. Tregenna, W. Flanagan, R. Maile, and V. Kendon, New J. Phys. **5**, 83 (2003).
[6] T. A. Brun, H. A. Carteret, and A. Ambainis, Phys. Rev. A **67**, 052317 (2003).
[7] A. Nayak and A. Vishwanath, e-print quant-ph/0010117.
[8] M. Hillery, J. Bergou, and E. Feldman, Phys. Rev. A **68**, 032314 (2003).
[9] O. Buerschaper and K. Burnett, e-print quant-ph/0406039.
[10] S. Venegas-Andraca, J. Ball, K. Burnett, and S. Bose, New J. Phys. **7**, 221 (2005).
[11] M. Dyer, A. Frieze, and R. Kannan, J. ACM **38**, 1 (1991).
[12] Q. Lv *et al.* (unpublished).
[13] B. C. Sanders, S. D. Bartlett, B. Tregenna, and P. L. Knight, Phys. Rev. A **67**, 042305 (2003).
[14] W. Dür, R. Raussendorf, V. M. Kendon, and H.-J. Briegel, Phys. Rev. A **66**, 052319 (2002).
[15] B. C. Travaglione and G. J. Milburn, Phys. Rev. A **65**, 032310 (2002).
[16] T. Schaetz (private communication).
[17] H. Jeong, M. Paternostro, and M. S. Kim, Phys. Rev. A **69**, 012310 (2004).
[18] R. C. Liu, B. Odom, Y. Yamamoto, and S. Tarucha, Nature (London) **391**, 263 (1998).
[19] Y. Omar, N. Paunković, S. Bose, and V. Vedral, Phys. Rev. A **65**, 062305 (2002); N. Paunković, Y. Omar, S. Bose, and V. Vedral, Phys. Rev. Lett. **88**, 187903 (2002); S. Bose, A. Ekert, Y. Omar, N. Paunković, and V. Vedral, Phys. Rev. A **68**, 052309 (2003).
[20] Y. Omar, Contemp. Phys. **46**, 437 (2005).

Optimal state discrimination using particle statistics

S. Bose,[1,2] A. Ekert,[3] Y. Omar,[4,5] N. Paunković,[4] and V. Vedral[6]
[1]*Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125, USA*
[2]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom*
[3]*Centre for Quantum Computation, DAMTP, University of Cambridge, Cambridge CB3 0WA, United Kingdom*
[4]*Centre for Quantum Computation, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*
[5]*Centro de Física de Plasmas, Instituto Superior Técnico, P-1049-001 Lisbon, Portugal*
[6]*Optics Section, The Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom*
(Received 24 April 2003; published 14 November 2003)

We present an application of particle statistics to the problem of optimal ambiguous discrimination of quantum states. The states to be discriminated are encoded in the internal degrees of freedom of identical particles, and we use the bunching and antibunching of the external degrees of freedom to discriminate between various internal states. We show that we can achieve the optimal single-shot discrimination probability using only the effects of particle statistics. We discuss interesting applications of our method to detecting entanglement and purifying mixed states. Our scheme can easily be implemented with the current technology.

DOI: 10.1103/PhysRevA.68.052309 PACS number(s): 03.67.Mn, 03.65.Wj, 05.30.−d

Recently, there has been an emerging interest in the use of particle statistics (both bosonic and fermionic) for quantum information processing [1,2]. In fact, it was shown that a useful task such as entanglement concentration could be accomplished, even if nonoptimally, using only the effects of quantum statistics, without the need for any other interactions [2]. The above investigations differ significantly from some previous suggestions, where either anyonic statistics [3] or the effects of electronic statistics in conjunction with other interactions [4,5] were used for quantum information processing. Schemes using only particle statistics [1,2] would be very useful for tasks implemented with identical particles that interact very weakly or not at all with each other, such as photons or neutrons. This weak interaction can be beneficial for information processing as it may reduce the unwanted coupling to the environment. Such schemes are also extremely general in the sense of being independent of the actual particle species. It is, however, not known whether such schemes can accomplish quantum information processing *efficiently*. Here we present a particular quantum information processing task involving two qubits and show that it can be performed *optimally* using *only* quantum statistics. Moreover, we point out how the task of discriminating quantum states can be applied to detecting entanglement and purifying mixed states. We also discuss how to generalize these tasks to N qubits and argue that quantum statistics could be used to perform even this generalized task optimally. While the two qubits and other small N qubits versions of our protocol can be tested with photons, electrons, neutrons, or atoms, the large number of qubits versions could have interesting implementations in optical lattices [6,7].

One of the striking aspects of quantum mechanics is that it is not possible to perfectly discriminate between two states unless they are orthogonal. Suppose someone prepares two qubits encoded in the internal degrees of freedom of two identical particles—say, in the spin of two electrons or the polarization of two photons—in one of the following two possible states: (1) spins aligned (parallel) and pointing in an arbitrary direction; (2) spins antialigned and pointing in an arbitrary direction.

We will assume that the two states are equally likely for the sake of simplicity, but all the presented results will be valid for any *a priori* distribution. Note also that, unless otherwise stated, whenever we mention spin we will actually refer to any two-dimensional internal degree of freedom, be it for fermions or for bosons. As we do not have any knowledge of the direction of alignment in both of the above cases, the overall states are mixed. They are described, respectively, by the following density operators:

$$\rho_2=\frac{1}{4\pi}\int d\Omega|\Omega\rangle\langle\Omega|\otimes|\Omega\rangle\langle\Omega| \tag{1}$$

and

$$\sigma_2=\frac{1}{4\pi}\int d\Omega|\Omega\rangle\langle\Omega|\otimes|\Omega^{\perp}\rangle\langle\Omega^{\perp}|. \tag{2}$$

Here, the subscripts indicate that we are considering two particles. The ket $|\Omega\rangle$ represents the spin-up state along the axis defined by the angle Ω, while $|\Omega^{\perp}\rangle$ is the orthogonal state spin down. Thus, ρ_2 represents an equal mixture of spins aligned along an arbitrary axis in space, while σ_2 is the equal mixture of antialigned spins of two spin-$\frac{1}{2}$ particles. It is impossible to discriminate between these states perfectly because they are not orthogonal.

Optimal results are known for the discrimination of any two given quantum states η and η' [8]. The maximal probability of ambiguously discriminating between two *a priori* equally likely quantum states in a single-shot measurement is given by the Helstrom formula:

$$P_H(\eta,\eta')=\tfrac{1}{2}+\tfrac{1}{4}\mathrm{Tr}|\eta-\eta'|. \tag{3}$$

We now present a procedure for discriminating between ρ_2 and σ_2, both for fermions and bosons, based only on the effects of particle statistics. To use these effects we interfere particles at a beam splitter. Here we use *beam splitter* in a generic sense, referring not only to the common optical element (partially silvered mirror) used with photons, but also

 ©2003 The American Physical Society

to any device presenting an analog behavior for other kinds of particles, as already suggested for electrons [9]. In order to distinguish the states we rely on path measurements that discriminate between bunching and antibunching, and that are performed on our particles after letting them pass simultaneously through a 50-50 beam splitter. Note that in such balanced beam splitters two indistinguishable particles will always bunch if they are bosons [10], and always antibunch if they are fermions [9] (see also Ref. [11]).

For fermions, our guess in the case of the antibunching result is that the input state was ρ_2, while in the case of bunching is that it was σ_2. The probability of success of our procedure is then

$$P_{BS}(\rho_2,\sigma_2)=\tfrac{3}{4}. \tag{4}$$

This probability can easily be calculated by noticing that the only case for which our guess could be incorrect is when we have the antibunching result. Then, we conclude that the input state was ρ_2, while it could have actually been σ_2. On the other hand, in the case of the bunching result we know for sure that the input state was σ_2, since—according to the Pauli exclusion principle—two particles with aligned spins cannot end up in the same output arm of the beam splitter. When our input state is σ_2, the antibunching happens with probability 1/2, giving in total a probability of incorrect inference of 1/4. In the case of bosons, our protocol is exactly the opposite of the fermionic one, but yields precisely the same efficiency. This time the antibunching results stand for σ_2, whereas the bunching ones stand for ρ_2, but the probability of success coincides with Eq. (4).

Interestingly, the Helstrom formula gives the same result for the maximal probability of discriminating between these two states:

$$P_H(\rho_2,\sigma_2)=\tfrac{3}{4}. \tag{5}$$

We can thus conclude that our procedure is optimal for both fermions and bosons.

Helstrom's probability can still be achieved in other cases using the effects of quantum statistics. We now introduce a case of special interest that can be applied to other tasks such as entanglement detection and state purification, as will be shown later in the paper. Suppose that we have to discriminate between the following two states: (1) spins aligned (parallel) and pointing in an arbitrary direction [the same as in Eq. (1)]; (2) each spin in the maximally mixed state.

The latter state is represented by the following operator:

$$\tau_2=\tfrac{1}{4}I^{\otimes 2}=\tfrac{1}{4}(|0\rangle\langle 0|+|1\rangle\langle 1|)\otimes(|0\rangle\langle 0|+|1\rangle\langle 1|), \tag{6}$$

where $|0\rangle$ and $|1\rangle$ are any two orthogonal spin states.

The strategy now is exactly the same as before. It relies on the fact that if the state is ρ_2 the particles can, due to their indistinguishability, give only one result (antibunching in the case of fermions or bunching in the case of bosons), while if the input state is τ_2 both results are possible. This time, the probability of success is

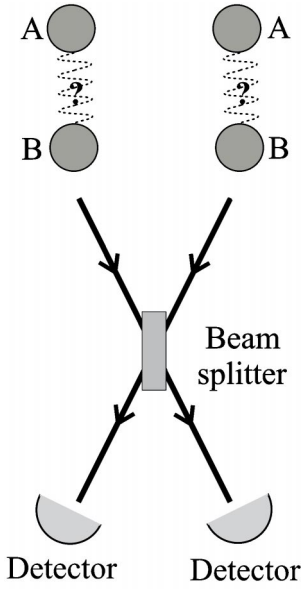


FIG. 1. This figure represents the setup for our entanglement detection scheme. We have two equal pairs of identical particles and consider their internal degrees of freedom. Both pairs are in the same pure state: either separable or maximally entangled. We take the same particle (for instance, B) from each pair and interfere them at a 50-50 beam splitter. If the states are disentangled then the particles are indistinguishable and, depending on the statistics, will either only bunch or only antibunch. Otherwise, if the states are maximally entangled, the particles are maximally mixed, meaning that they can be (probabilistically) distinguished and hence the statistics does not influence their behavior. This entanglement detection is a particular instance of our state discrimination scheme discussed in the paper.

$$P_{BS}(\rho_2,\tau_2)=\tfrac{5}{8}, \tag{7}$$

and this coincides with the Helstrom result

$$P_H(\rho_2,\tau_2)=\frac{5}{8}, \tag{8}$$

so our procedure is optimal in this case too. The calculations leading to the above results are analogous to the ones in the previous discrimination case.

One of the most interesting applications of our approach can be found in the detection of entanglement. We illustrate this in the case of pure states of two particles. Suppose, for example, that we are to discriminate any maximally entangled state of two qubits from any disentangled (product) state. As before, the qubits are supported by the internal degrees of freedom of two identical particles (labeled, say, A and B). Suppose in addition that we are given two identical copies of the state. In order to detect entanglement, we take the same particle (either A or B) from each pair and interfere them at a beam splitter, as shown in Fig. 1. The crux of the argument is that if the state is entangled, then the reduced states of these particles will be maximally mixed as in the state τ_2. On the other hand, if the state is separable, then the two interfering particles are in the state ρ_2. This is the same as in our discrimination procedure above, and so there is a

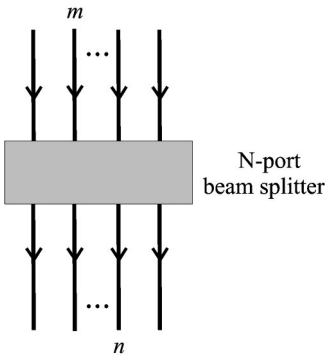


FIG. 2. This diagram represents a multiport beam splitter with N inputs and N outputs. The overall output state depends not only on the input, but also on the statistics (either fermionic or bosonic) of the identical particles involved. We have labeled two arbitrary ports, the mth input port and the nth output port.

probability of 5/8 to detect entanglement. This example can be generalized to other entangled pure states, and, more interestingly, to more particles. We note that there is a close analogy between this method and our entanglement concentration scheme in Ref. [2]. In particular, if the two states were less than maximally entangled, then by detecting entanglement we would actually also amplify it (see Ref. [2] for more details).

Another interesting application is in mixed state purification, as in Ref. [12]. Suppose that we start with two qubits, each in some mixed state. We would like not only to make the state of these qubits purer (in the sense of having lower linear entropy), but also to preserve their original direction in the Bloch sphere. The optimal way of doing so (as proven in Ref. [12]) is to project the joint state onto the symmetric subspace, in which case the resulting mixed state is purer and yet preserves the original direction. If the projection is unsuccessful, the qubits are thrown away. This is exactly the same as our probabilistic discrimination with a beam splitter.

We would now like to investigate the generalization of the above results to N particles. For this, we use a generalized (N-port) balanced beam splitter, as shown in Fig. 2, which acts only on the spatial degrees of freedom of the input particles. This action is given by a unitary matrix U_N, with elements

$$u_{mn}=\frac{1}{\sqrt{N}}e^{i(2\pi/N)(m-1)(n-1)}. \tag{9}$$

(Note that there exist alternative descriptions of balanced multiport beam splitters [13].) The square of the norm of each element in the matrix represents the probability that the particle in the mth input arm of the beam splitter ends up in the nth output arm. Since all these elements have norm $1/\sqrt{N}$, we have a representation of a balanced N-port beam splitter.

Now the aim is to discriminate between the N-particle generalizations of ρ_2 and τ_2. Those states are given by

$$\rho_N=\frac{1}{4\pi}\int d\Omega(|\Omega\rangle\langle\Omega|)^{\otimes N} \tag{10}$$

and

$$\tau_N=\frac{1}{2^N}(|0\rangle\langle 0|+|1\rangle\langle 1|)^{\otimes N}. \tag{11}$$

To calculate the Helstrom probability, one has to diagonalize the matrix $\rho_N-\tau_N$. This turns out to be straightforward once we notice that ρ_N can also be represented as an equal mixture of all possible symmetric states within the basis $\{|S_i\rangle: i=1,\ldots,N+1=|S|\}$ of N qubits. Then, we have

$$\rho_N=\frac{1}{|S|}\sum_{i=1}^{|S|}|S_i\rangle\langle S_i|. \tag{12}$$

It is now easy to calculate the Helstrom formula by expanding ρ_N in a basis consisting of the union of a basis of the symmetric subspace and a basis of its orthogonal complement. The result is

$$P_H(\rho_N,\tau_N)=1-\frac{(N+1)}{2^{(N+1)}}. \tag{13}$$

Alternatively, we can calculate the average probability of success to distinguish states ρ_N and τ_N using the following expression:

$$P_H(\rho_N,\tau_N)=\tfrac{1}{2}1+\tfrac{1}{2}p. \tag{14}$$

Here, the 1/2 factors refer to the fact that the two states are prepared with equal probability. The term 1 comes from the fact that the state ρ_N, supported on the symmetric subspace (of dimension $d_S=N+1$), is always identified reliably as such, and p is the probability of identifying the other state τ_N. Since τ_N is maximally mixed, it is uniformly distributed over the whole space of N qubits (of dimension $d=2^N$). In this case $p=(1-d_S)/d$, which after substitution in Eq. (14) gives Eq. (13) right away.

The problem to apply our discrimination scheme to N particles is that it becomes exponentially hard to calculate $P_{BS}(\rho_N,\tau_N)$ as N increases. Moreover, it is not clear which inference strategy should be followed. For fermions, the natural generalization would be to associate the antibunching results with ρ_N and the others with τ_N. For bosons, on the other hand, a more subtle strategy may be needed. This is because, loosely speaking, for bosons there is no clear analog of the Pauli exclusion principle. Furthermore, even without having the complete calculations for $N>2$, we would like to emphasize the remarkable fact that the Helstrom probability $P_H(\rho_N,\tau_N)$ is equal to the probability of success of a fermionic beam splitter strategy described above, if calculated under the assumption that the particles are classical (i.e., always distinguishable by some arbitrary label), but obey a constraint equivalent to the Pauli exclusion principle (not allowing more than two particles in the same internal state to share the same output arm of the beam splitter). The overall probability is then calculated by summing up the probabilities of

all the possible outcomes rather than the amplitudes, as it would be done in the quantum case. However, in the case of three fermions ($N=3$), we have performed the full quantum calculations (i.e., taking properly into account the effects of statistics) for a three-port balanced beam splitter and obtained

$$P_{BS}(\rho_3,\tau_3)=\tfrac{3}{4}, \tag{15}$$

which is equal to $P_H(\rho_3,\tau_3)$. We believe this result of obtaining the optimal discrimination probability using the effects of particle statistics (in multiports) can be generalized to an arbitrary N, both for fermions and bosons, and we continue research in this direction. For now this remains a conjecture. In an optical lattice, with one particle in each lattice site, a multiport beam splitter could probably be simulated by dissolving N potential wells and then creating a new set of N wells [7]. Of course, if the particles interact, then the effective beam splitter will be modified, and here we only point out the plausibility of creating multiport beam splitters (or multiparticle interference) in an optical lattice.

In this paper we have shown that it is possible to perform an optimal quantum information processing task using only the effects of particle statistics. In particular, we have presented a strategy for discriminating between two nonorthogonal states of two qubits (encoded in the internal degrees of freedom of identical particles) using beam splitters. We have considered two discrimination scenarios and in each of them our strategy differs (symmetrically) between fermions and bosons, but offers the same efficiency. We also pointed out how our discrimination scheme can be applied to detecting entanglement and purifying mixed states. In addition, we have calculated the Helstrom probability for N qubits in one of our discrimination scenarios. We have shown by explicit calculation that this probability can be achieved in a fermionic three-port beam splitter strategy and that it is the same as the fermionic strategy for general N if the fermions are considered as classical particles that obey the Pauli exclusion principle as the only additional constraint. An advantage of our method is that it can also be easily implemented with the current technology.

Our work suggests a number of interesting research directions. One problem is to prove the optimality of the beam splitter strategy in the case of N qubits and its application to multiparty entanglement detection. This, we hope, will answer the question of weather the symmetry between fermions and bosons in our strategy will be preserved for a generalized beam splitter. It may also lead to a simple and physically intuitive selection principle governing bosonic behavior. Another possible direction is to classify all the pairs of states that can be optimally discriminated with our scheme. Finally, our results suggest that it would be worth to further explore the role of particle statistics in quantum information tasks, in particular in efficient quantum computation. After the conclusion of this paper, several interesting related works were pointed out to us [14–16].

The authors would like to thank G. Castagnoli for many useful discussions on the link between computational complexity and particle statistics. His ideas have stimulated much of the work in this paper. We would also like to thank F. A. Bovino and P. Varisco for useful discussions on the role of particle statistics in quantum information processing. S.B. acknowledges support from the NSF under Grant No. EIA-00860368. Y.O. acknowledges support from Fundação para a Ciência e a Tecnologia from Portugal and the Institute for Quantum Information at Caltech for their hospitality. N.P. and V.V. acknowledge Elsag SpA for financial support and Centro de Física de Plasmas at IST in Lisbon for their hospitality. V.V. acknowledges support from Hewlett-Packard company, EPSRC and the European Union projects EQUIP, and TOPQIP. This work was partly supported by the QUIPROCONE collaboration Grant No. 044.

[1] Y. Omar, N. Paunković, S. Bose, and V. Vedral, Phys. Rev. A **65**, 062305 (2002).
[2] N. Paunković, Y. Omar, S. Bose, and V. Vedral, Phys. Rev. Lett. **88**, 187903 (2002).
[3] A.Yu. Kitaev, Ann. Phys. (N.Y.) **303**, 2-30 (2003).
[4] D. Loss and D.P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).
[5] A.T. Costa, Jr. and S. Bose, Phys. Rev. Lett. **87**, 277901 (2001).
[6] J. Pachos and P.L. Knight, e-print quant-ph/0301084.
[7] D. Jaksch *et al.*, Phys. Rev. Lett. **82**, 1975 (1999).
[8] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[9] R.C. Liu, B. Odom, Y. Yamamoto, and S. Tarucha, Nature (London) **391**, 263 (1998).
[10] C.K. Hong, Z.Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).
[11] R. Loudon, Phys. Rev. A **58**, 4904 (1998).
[12] J.I. Cirac, A.K. Ekert, and C. Macchiavello, Phys. Rev. Lett. **82**, 4344 (1999).
[13] I. Jex, S. Stenholm, and A. Zeilinger, Opt. Commun. **117**, 95 (1995); see also M. Żukowski, A. Zeilinger, and M.A. Horne, Phys. Rev. A **55**, 2564 (1997).
[14] S.M. Barnett, A. Chefles, and I. Jex, Phys. Lett. A **307**, 189 (2003).
[15] J. Fiurasek, M. Dusek, and R. Filip, Phys. Rev. Lett. **89**, 190401 (2002); M. Hendrych, M. Dusek, R. Filip, and J. Fiurasek, Phys. Lett. A **310**, 95 (2003).
[16] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Phys. Rev. Lett. **91**, 027901 (2003).

Entanglement Concentration Using Quantum Statistics

N. Paunković,[1] Y. Omar,[1] S. Bose,[1] and V. Vedral[2]
[1]*Centre for Quantum Computation, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, United Kingdom*
[2]*Optics Section, The Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom*
(Received 1 December 2001; published 18 April 2002)

We propose an entanglement concentration scheme which uses only the effects of quantum statistics of indistinguishable particles. This establishes the fact that useful quantum information processing can be accomplished by quantum statistics alone. Because of the basis independence of statistical effects, our protocol requires less knowledge of the initial state than most entanglement concentration schemes. Moreover, no explicit controlled operation is required at any stage.

DOI: 10.1103/PhysRevLett.88.187903 PACS numbers: 03.67.–a, 03.65.Ud, 05.30.–d

Quantum statistics has led to a number of interesting phenomena such as ferromagnetism, superconductivity, and superfluidity. However, its effects have never been used to process information. Recently, we have shown that quantum statistics can lead to an *effective* interaction between internal and external degrees of freedom of particles [1]. This was demonstrated in the context of entanglement transfer from spin to path without an explicit conditional operation between these degrees of freedom. Going even further in our exploration of the effects of quantum statistics in information theory, we present an entanglement concentration protocol [2,3] based on these effects. This establishes the fact that quantum statistics *alone* (without any other explicit interaction between the relevant degrees of freedom of the particles involved) is sufficient for useful quantum information processing.

The best performances in quantum communication and computation processes are normally achieved using a pair of pure maximally entangled systems (particles) shared between distant parties. But the inevitable influence of the environment during the distribution of the entangled pairs reduces the amount of shared entanglement. As entanglement cannot be increased by local operations and classical communication [4], the only option is to concentrate it locally from a larger to a smaller number of pairs [2,3,5–11]. In the case of mixed states we call these protocols *entanglement purification,* while for pure states we refer to them as *entanglement concentration* [3].

In a setting similar to previous entanglement concentration protocols, we consider the action of specific local operations on two pairs of entangled systems, driving them with some probability into one pair of more entangled systems.

However, unlike most of the previous schemes, we require our entangled systems to be composed of n particles (see Fig. 1 for the representation of a pair of these systems). We look at the entanglement in the internal degrees of freedom of the particles, such as the spin in the case of electrons (fermions) or the polarization in the case of photons (bosons), which have isomorphic Hilbert spaces. The initial pure state of each of our two pairs, distributed between two parties Alice (A) and Bob (B), is

$$|\phi\rangle^n \equiv \alpha|\underbrace{\uparrow\uparrow\cdots\uparrow}_{n}\rangle_A|\downarrow\downarrow\cdots\downarrow\rangle_B + \beta|\downarrow\downarrow\cdots\downarrow\rangle_A|\uparrow\uparrow\cdots\uparrow\rangle_B \tag{1}$$

with $|\alpha|^2 + |\beta|^2 = 1$ and where, for instance, we have

$$|\uparrow\uparrow\cdots\uparrow\rangle_A = \hat{a}^+_{A1\uparrow}\hat{a}^+_{A2\uparrow}\cdots\hat{a}^+_{An\uparrow}|0\rangle \equiv |A\tilde{\uparrow}\rangle^n. \tag{2}$$

Here $\hat{a}^+$ are creation operators that ensure the ordering of the multiparticle state, as the usual commutation and anticommutation rules apply for bosons and fermions, respectively. For the sake of compactness, we will rewrite Eq. (1) as

$$|\phi\rangle^n = (\alpha|A\tilde{\uparrow}\rangle^n|B\tilde{\downarrow}\rangle^n + \beta|A\tilde{\downarrow}\rangle^n|B\tilde{\uparrow}\rangle^n), \tag{3}$$

where the tilde over the arrow reminds us of the string of spins (or polarizations).

Let us now label one pair L and the other R (as in left and right, see Fig. 2). Then, our total initial state is

$$|\varphi\rangle^n \equiv |\phi\rangle^n_L \otimes |\phi\rangle^n_R, \tag{4}$$

where $|\phi\rangle^n_L$ is given by Eq. (1) written for pair L, and similarly for R.

Using Eq. (3), we can rewrite the total state (4) as

$$|\varphi\rangle^n = \alpha^2|A\tilde{\uparrow}\rangle^n_L|A\tilde{\uparrow}\rangle^n_R|B\tilde{\downarrow}\rangle^n_L|B\tilde{\downarrow}\rangle^n_R + \beta^2|A\tilde{\downarrow}\rangle^n_L|A\tilde{\downarrow}\rangle^n_R|B\tilde{\uparrow}\rangle^n_L|B\tilde{\uparrow}\rangle^n_R + \alpha\beta(|A\tilde{\uparrow}\rangle^n_L|A\tilde{\downarrow}\rangle^n_R|B\tilde{\downarrow}\rangle^n_L|B\tilde{\uparrow}\rangle^n_R + |A\tilde{\downarrow}\rangle^n_L|A\tilde{\uparrow}\rangle^n_R|B\tilde{\uparrow}\rangle^n_L|B\tilde{\downarrow}\rangle^n_R). \tag{5}$$

Our protocol consists of bringing each of Bob's particles into a 50/50 beam splitter (particles i from the left and the right systems go into beam splitter i, for $i = 1, 2, \ldots, n$) and then making a specific measurement on the output particles

 © 2002 The American Physical Society

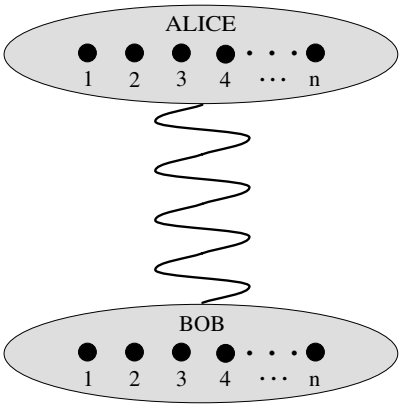


FIG. 1. This figure represents one of the two entangled pairs that we use for our protocol. Each pair is composed of two n-particle states entangled between Alice and Bob with at most 1 e-bit of entanglement.

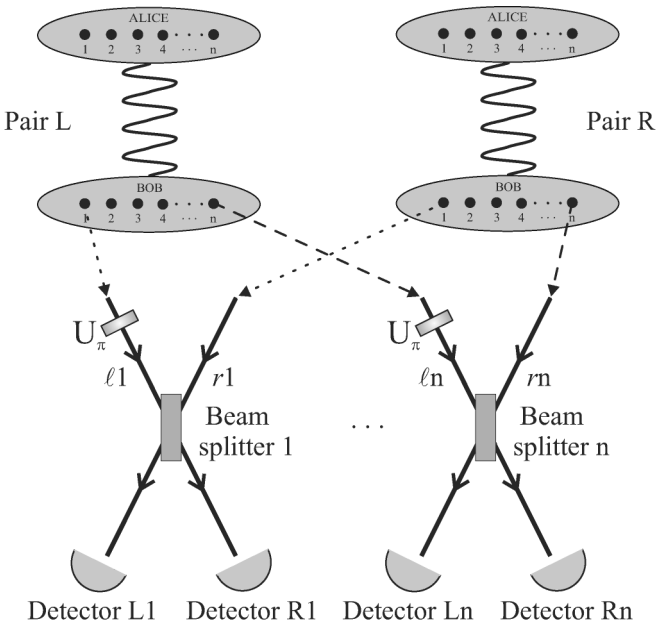


FIG. 2. This figure represents the setup for our entanglement concentration protocol using quantum statistics. Initially, Alice and Bob share two pairs of entangled systems, L and R. Each entangled pair is composed of two n-particle states. One of the parties, say Bob, can then convert these pairs into a more entangled pair by performing a set of local operations using only standard 50/50 beam splitters, path measurements, and one-way classical communication with Alice.

using detectors Li and Ri (see Fig. 2). Before the particles from the left fall into the beam splitter, we apply to each of them a specific unitary transformation U_π^i, that flips their spin or polarization. This is done in the arm ℓi of each beam splitter and the transformation is defined for every i:

$$U_\pi^i : \begin{cases} |B\uparrow\rangle_{\ell i} \rightarrow |B\downarrow\rangle_{\ell i}\,, \\ |B\downarrow\rangle_{\ell i} \rightarrow |B\uparrow\rangle_{\ell i}\,. \end{cases} \tag{6}$$

This flipping is applied in order to align some spins (or polarizations) in appropriate terms so that we can exploit the statistical effects later on.

The total state then becomes

$$\begin{aligned} |\varphi_0\rangle^n &\equiv (U_\pi^1 \otimes U_\pi^2 \otimes \cdots \otimes U_\pi^n)\,|\varphi\rangle^n \\ &= \alpha^2 |A\tilde{\uparrow}\rangle_L^n |A\tilde{\uparrow}\rangle_R^n |B\tilde{\uparrow}\rangle_L^n |B\tilde{\downarrow}\rangle_R^n + \beta^2 |A\tilde{\downarrow}\rangle_L^n |A\tilde{\downarrow}\rangle_R^n |B\tilde{\downarrow}\rangle_L^n |B\tilde{\uparrow}\rangle_R^n + \alpha\beta(|A\tilde{\uparrow}\rangle_L^n |A\tilde{\downarrow}\rangle_R^n |B\tilde{\uparrow}\rangle_L^n |B\tilde{\uparrow}\rangle_R^n + |A\tilde{\downarrow}\rangle_L^n |A\tilde{\uparrow}\rangle_R^n |B\tilde{\downarrow}\rangle_L^n |B\tilde{\downarrow}\rangle_R^n)\,. \end{aligned} \tag{7}$$

First, we present the entanglement concentration protocol for fermions (e.g., electrons) and then the counterpart protocol for bosons (e.g., photons).

Fermions.—After the particles pass through the beam splitter, Bob performs a path measurement on the first pair using detectors $L1$ and $R1$ (see Fig. 2). For simplicity of presentation, we first assume that these detectors do not absorb the particles and do not disturb their internal degrees of freedom (such detectors are indeed available for electrons and atoms and can have interesting applications in deterministically entangling independent particles using indistinguishability and feedback [12]). However, as we will show later on, this is not a necessary requirement for success of our protocol. If the particles (which we assumed to be fermions) bunch, we discard them. Otherwise, in the case of antibunching, we get the pure state

$$\begin{aligned} |\varphi_1\rangle^n = N_1 \Big[\frac{1}{\sqrt{2}} &(\alpha^2 |A\tilde{\uparrow}\rangle_L^n |A\tilde{\uparrow}\rangle_R^n |B\,\text{triplet}\rangle_{L1R1} |B\tilde{\uparrow}\rangle_L^{n-1} |B\tilde{\downarrow}\rangle_R^{n-1} + \beta^2 |A\tilde{\downarrow}\rangle_L^n |A\tilde{\downarrow}\rangle_R^n |B\,\text{triplet}\rangle_{L1R1} |B\tilde{\downarrow}\rangle_L^{n-1} |B\tilde{\uparrow}\rangle_R^{n-1}) \\ &+ \alpha\beta(|A\tilde{\uparrow}\rangle_L^n |A\tilde{\downarrow}\rangle_R^n |B\tilde{\uparrow}\rangle_L^n |B\tilde{\uparrow}\rangle_R^n + |A\tilde{\downarrow}\rangle_L^n |A\tilde{\uparrow}\rangle_R^n |B\tilde{\downarrow}\rangle_L^n |B\tilde{\downarrow}\rangle_R^n) \Big], \end{aligned} \tag{8}$$

where

$$N_1 = \left[\frac{|\alpha|^4}{2} + \frac{|\beta|^4}{2} + 2|\alpha\beta|^2 \right]^{-(1/2)}, \tag{9}$$

and the state

$$|B\,\text{triplet}\rangle_{L1R1} \equiv \frac{1}{\sqrt{2}} (|B\uparrow\rangle_{L1} |B\downarrow\rangle_{R1} + |B\downarrow\rangle_{L1} |B\uparrow\rangle_{R1}) \tag{10}$$

is a triplet state, the antibunching result of Bob's measurement with detectors $L1$ and $R1$ on the first pair of particles.

The probability p_1 of having antibunching and thus getting state $|\varphi_1\rangle^n$ (in an ideal measurement) is

$$p_1 = \frac{|\alpha|^4}{2} + \frac{|\beta|^4}{2} + 2|\alpha\beta|^2 = N_1^{-2}, \tag{11}$$

because the first two terms have a probability $\frac{1}{2}$ of antibunching, while the other two have probability 1 (we normalize these probabilities by $|\alpha|^2$, $|\beta|^2$, and $|\alpha\beta|^2$, the probabilities of obtaining the corresponding results). Thus, the state $|\varphi_1\rangle$ differs from $|\varphi_0\rangle$ in the first two terms (we "normalize" their amplitudes α^2 and β^2 by the factor $\frac{1}{\sqrt{2}}$).

After obtaining the antibunching result, we can now let the second pair of Bob's particles ($\ell 2$ and $r2$) pass through another 50/50 beam splitter and perform the same measurement, discarding again the bunching results.

After repeating the same procedure n times, we get the state

$$|\varphi_n\rangle^n = N_n\Big[\frac{\alpha^2}{2^{n/2}}\,|A\tilde{\uparrow}\rangle_L^n|A\tilde{\uparrow}\rangle_R^n|B\,\text{triplet}\rangle^{\otimes n} + \frac{\beta^2}{2^{n/2}}\,|A\tilde{\downarrow}\rangle_L^n|A\tilde{\downarrow}\rangle_R^n|B\,\text{triplet}\rangle^{\otimes n} + \alpha\beta(|A\tilde{\uparrow}\rangle_L^n|A\tilde{\downarrow}\rangle_R^n|B\tilde{\uparrow}\rangle_L^n|B\tilde{\uparrow}\rangle_R^n + |A\tilde{\downarrow}\rangle_L^n|A\tilde{\uparrow}\rangle_R^n|B\tilde{\downarrow}\rangle_L^n|B\tilde{\downarrow}\rangle_R^n)\Big], \tag{12}$$

with

$$N_n = \left[\frac{|\alpha|^4}{2^n} + \frac{|\beta|^4}{2^n} + 2|\alpha\beta|^2\right]^{-(1/2)}. \tag{13}$$

The probability of getting $|\varphi_n\rangle$ from $|\varphi_{n-1}\rangle$ is

$$p_n = \left(\frac{N_{n-1}}{N_n}\right)^2. \tag{14}$$

So, the total probability for getting $|\varphi_n\rangle^n$ from $|\varphi_0\rangle^n$ is

$$\tilde{p}_n = \prod_{i=1}^{n} p_i = N_n^{-2} \xrightarrow{(n\to\infty)} 2|\alpha\beta|^2 \equiv \tilde{p}\,. \tag{15}$$

After applying the same procedure "infinitely many times" (to an "infinitely large state"), we obtain the state

$$|\varphi_\infty\rangle \equiv \lim_{n\to\infty}|\varphi_n\rangle^n = \lim_{n\to\infty}\frac{1}{\sqrt{2}}\,(|A\tilde{\uparrow}\rangle_L^n|A\tilde{\downarrow}\rangle_R^n|B\tilde{\uparrow}\rangle_L^n|B\tilde{\uparrow}\rangle_R^n + |A\tilde{\downarrow}\rangle_L^n|A\tilde{\uparrow}\rangle_R^n|B\tilde{\downarrow}\rangle_L^n|B\tilde{\downarrow}\rangle_R^n)\,. \tag{16}$$

The total probability is then

$$p = \frac{\tilde{p}}{2} = |\alpha\beta|^2. \tag{17}$$

Now note that even if we had, instead of nonabsorbing, any kind of path detectors, our protocol would still work assuming we perform a path measurement on $n-1$ particles. We would then have absorbed all these $n-1$ particles, and their state would be replaced by the vacuum state $|0\rangle$. In the $n\to\infty$ limit this will factorize out to give the final maximally entangled state:

$$|\varphi_\infty\rangle = \lim_{n\to\infty}\frac{1}{\sqrt{2}}\,(|A\tilde{\uparrow}\rangle_L^n|A\tilde{\downarrow}\rangle_R^n|B\uparrow\rangle_L|B\uparrow\rangle_R + |A\tilde{\downarrow}\rangle_L^n|A\tilde{\uparrow}\rangle_R^n|B\downarrow\rangle_L|B\downarrow\rangle_R)\,. \tag{18}$$

The probability p given by Eq. (17) provides us with a measure of efficiency of our entanglement concentration protocol. It is the amount of entanglement in e-bits that we can extract from a single pair of entangled particles in the initial state $|\phi\rangle^n$ given by Eq. (3), in the limiting case:

$$E^\infty(|\phi\rangle^n) = p, \qquad (n\to\infty)\,. \tag{19}$$

Bosons.— In this case, the procedure is almost the same as for fermions, but this time we discard the antibunching results and keep the bunching ones.

After Bob measures the first pair of particles, we obtain a state similar to (8), but where the particles will now be either both in $L1$ or either both in $R1$. In other words, instead of the triplet state $|B\,\text{triplet}\rangle_{L1R1}$, we have the term $\frac{1}{\sqrt{2}}(|B\uparrow\rangle_{L1}|B\downarrow\rangle_{L1} + |B\uparrow\rangle_{R1}|B\downarrow\rangle_{R1})$, while instead of $|B\tilde{\uparrow}\rangle_L^n|B\tilde{\uparrow}\rangle_R^n$ we have $\frac{1}{\sqrt{2}}(|B\tilde{\uparrow}\rangle_{L1}|B\tilde{\uparrow}\rangle_{L1} + |B\tilde{\uparrow}\rangle_{R1}|B\tilde{\uparrow}\rangle_{R1}) \otimes |B\tilde{\uparrow}\rangle_L^{n-1}|B\tilde{\uparrow}\rangle_R^{n-1}$, and similarly for the last term in (8).

The probability of getting this state is again given by (11), $p_1 = N_1^{-2}$.

By reversing the selection of the path measurements performed by Bob (i.e., by now discarding the antibunching instead of the bunching results), we have established a symmetry between the bosonic and the fermionic protocols. The general state (12) and its probability (14) are in fact the "same" (isomorphic), as well as the total probability (17), $p = |\alpha\beta|^2$, and the efficiency (19):

$$E^\infty(|\phi\rangle^n) = p, \qquad (n\to\infty)\,. \tag{20}$$

Note that for both protocols (for fermions and for bosons) the unitary transformation $U_\pi \equiv \bigotimes_{i=1}^n U_\pi^i$ is crucial. Since states $|\varphi\rangle^n$ and $|\varphi_0\rangle^n = U_\pi|\varphi\rangle^n$, given by Eqs. (5) and (7), respectively, are isomorphic, one might expect the results to be the same regardless of U_π being applied or not. This turns out to be wrong, as applying the protocol directly to $|\varphi\rangle^n$ will yield the following state:

$$|\varphi_\infty\rangle \equiv \lim_{n\to\infty}|\varphi_n\rangle = \lim_{n\to\infty}\frac{1}{\sqrt{|\alpha|^4 + |\beta|^4}}\,(\alpha^2|A\tilde{\uparrow}\rangle_L^n|A\tilde{\uparrow}\rangle_R^n|B\tilde{\downarrow}\rangle_L^n|B\tilde{\downarrow}\rangle_R^n + \beta^2|A\tilde{\downarrow}\rangle_L^n|A\tilde{\downarrow}\rangle_R^n|B\tilde{\uparrow}\rangle_L^n|B\tilde{\uparrow}\rangle_R^n)\,, \tag{21}$$

in general less entangled than the initial one.

In the protocols we have presented here, local operations are performed on one side only (Bob). Classical communication comes about only once, when we have one way communication from Bob to Alice at the end of the whole scheme. Of course, we can slightly change our protocols by allowing Alice to apply the same procedure on her side. This would require two-way communication and would have half the time complexity of the protocol, but the amount of entanglement distilled would be the same. Also, note that both Alice and Bob could perform the operations on the $n-1$ particles, each one on their side, using either one beam splitter sequentially or $n-1$ beam

splitters in parallel. The latter case has a lower time complexity than the former, but requires more resources (higher space complexity).

The efficiency $|\alpha\beta|^2$ of our protocol happens to be lower than both the efficiency $2|\alpha|^2$ of the procrustean method and the efficiency $-|\alpha|^2\,\mathrm{Log}|\alpha|^2 - |\beta|^2\,\mathrm{Log}|\beta|^2$ of the standard asymptotic entanglement concentration procedure [3]. However, our target is *not* to propose a more efficient *alternative* entanglement concentration, but to explicitly demonstrate that it is *possible* to do entanglement concentration by using *only* the effects of particle statistics, without resort to an explicit controlled operation between spins or between spin and path. In all physical implementations of entanglement concentration or purification protocols to date [10,11], a polarization beam splitter or a polarization dependent filter (which accomplish a controlled operation on path conditional on spin) have been used.

It is important to note here that our mechanism will work even if the basis states $|\uparrow\rangle$ and $|\downarrow\rangle$, were rotated in a plane. This is because the only unitary rotation we use is U_π, which will flip all spins in one plane. The rest of the protocol uses particle statistics, which is basis independent. On the other hand, for standard entanglement concentration protocols [2,3], it is necessary to know the basis.

In this paper we have presented an entanglement concentration scheme which uses *only* the effects of quantum statistics. Although the efficiency of the protocol is the same for both fermions and bosons, the protocol itself is slightly different depending on the nature of the particles. This brings forth a fundamental difference between these two types of particles in terms of their information processing power. Recent experiments such as [13,14] suggest that it would be possible to test our results in the near future. It would also be interesting to investigate the possibility of entanglement distillation of mixed states using *solely* statistical effects. Future work will comprise investigating more extensive quantum information processing procedures.

N. P. thanks Elsag S.p.A. for financial support. Y. O. acknowledges support from Fundação para a Ciência e a Tecnologia from Portugal. V. V. acknowledges support from Hewlett-Packard company, EPSRC, and the European Union project EQUIP.

[1] Y. Omar, N. Paunković, S. Bose, and V. Vedral, quant-ph/0105120 [Phys. Rev. A (to be published)].
[2] N. Gisin, Phys. Lett. A **210**, 151 (1996).
[3] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[4] M. Plenio and V. Vedral, Contemp. Phys. **39**, 431 (1998).
[5] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[6] D. Deutsch *et al.,* Phys. Rev. Lett. **77**, 2818 (1996).
[7] S. Bose, V. Vedral, and P. L. Knight, Phys. Rev. A **60**, 194 (1999); L. Hardy and D. D. Song, Phys. Rev. A **62**, 052315 (2000).
[8] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **78**, 574 (1997).
[9] R. T. Thew and W. J. Munro, Phys. Rev. A **63**, 030302 (2001).
[10] P. G. Kwiat, S. Barraza-Lopez, A. Stefanov, and N. Gisin, Nature (London) **409**, 1014 (2001).
[11] J. W. Pan, C. Simon, Č. Brukner, and A. Zeilinger, Nature (London) **410**, 1067 (2001).
[12] S. Bose and D. Home, Phys. Rev. Lett. **88**, 050401 (2002).
[13] T. J. Herzog, P. G. Kwiat, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **75**, 3034 (1995).
[14] R. C. Liu, B. Odom, Y. Yamamoto, and S. Tarucha, Nature (London) **391**, 263 (1998).

Spin-space entanglement transfer and quantum statistics

Y. Omar,[1] N. Paunković,[1] S. Bose,[1] and V. Vedral[2]
[1]*Centre for Quantum Computation, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, United Kingdom*
[2]*Optics Section, The Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom*
(Received 9 August 2001; published 30 May 2002)

Both the topics of entanglement and particle statistics have aroused enormous research interest since the advent of quantum mechanics. Using two pairs of entangled particles we show that indistinguishability enforces a transfer of entanglement from the internal to the spatial degrees of freedom without any interaction between these degrees of freedom. Moreover, subensembles selected by local measurements of the path will, in general, have different amounts of entanglement in the internal degrees of freedom depending on the statistics (either fermionic or bosonic) of the particles involved.

DOI: 10.1103/PhysRevA.65.062305 PACS number(s): 03.67.−a, 03.65.Ud, 05.30.−d

Since the advent of quantum mechanics, entanglement has been identified as one of its most peculiar features [1–3]. This ''excess correlation'' has recently become an important resource in quantum information processing [4]. Entanglement is believed to be at the root of the speedup of quantum computers over their classical counterparts [5], and it also leads to an unconditionally secure quantum cryptographic key exchange [6]. Another fundamental aspect of quantum physics, somewhat neglected in the field of quantum information, is the distinction between two different types of particles, fermions and bosons, manifested through particle statistics (although see [7] and for fermions see [8–10]). There are at first sight two seemingly ''conflicting'' views regarding the role of indistinguishability and particle statistics in quantum information processing. On the one hand, these two notions appear to combine to offer ''natural'' entanglement through forcing the use of symmetrized and antisymmetrized states (for bosons and fermions, respectively), and as we mentioned before, entanglement is generally an advantage for quantum information processing (although see [11]). On the other hand, indistinguishability prevents us from addressing the particles separately that seems to be a disadvantage in information processing. In this paper we analyze the role of indistinguishability and particle statistics in a simple information processing scenario.

Consider the following situation. Suppose that we have two pairs of qubits (quantum two-level systems), each pair maximally entangled in some internal degree of freedom. If the particles carrying the qubits are of the same type—say bosons—but distinguishable as a result of spatial separation, then we have two units of entanglement (ebits) in total. All of this entanglement is in the internal degrees of freedom. If we now consider bringing the particles close together and then separating them again, without the internal degrees of freedom ever interacting with the spatial ones, we should expect the whole entanglement to remain in the internal degrees of freedom. Surprisingly, as we demonstrate in this paper, a fraction of the initial entanglement is transferred to the path degrees of freedom of the particles. The fascinating implication is that the transfer of entanglement is imposed by particle indistinguishability and does not involve any controlled operation between the internal and external degrees of freedom (i.e., spin-path interaction), in contrast with the standard entanglement swapping scheme [12]. The prevalent setting for local manipulations of entanglement in quantum information processing either involves explicit interactions between the internal degrees of freedom of two particles, or an interaction of the internal degrees of freedom with some apparatus. Here we introduce a completely different setting in which particle paths are locally mixed without any interaction of the internal degrees of freedom with anything else.

Now we turn to describe the exact details of our thought experiment. Imagine the following setup, described in Fig. 1. We have two pairs of identical particles, each pair being maximally entangled in some internal degree of freedom, e.g., the spin or polarization. In our case, we consider systems with spin one-half, or isomorphic to it. We assume that our setup is symmetrical both horizontally and vertically, where the dotted lines in Fig. 1 show the axis of symmetry.

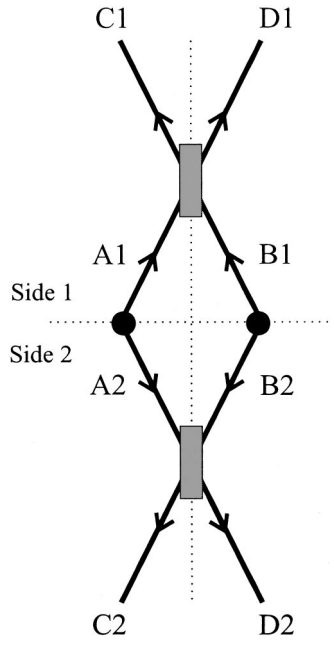


FIG. 1. This figure presents our setup for spin-space entanglement transfer. Each black circle represents a source of a pair of particles maximally entangled in the internal degrees of freedom (not explicitly shown in the figure). The rectangles represent beam splitters.

 ©2002 The American Physical Society

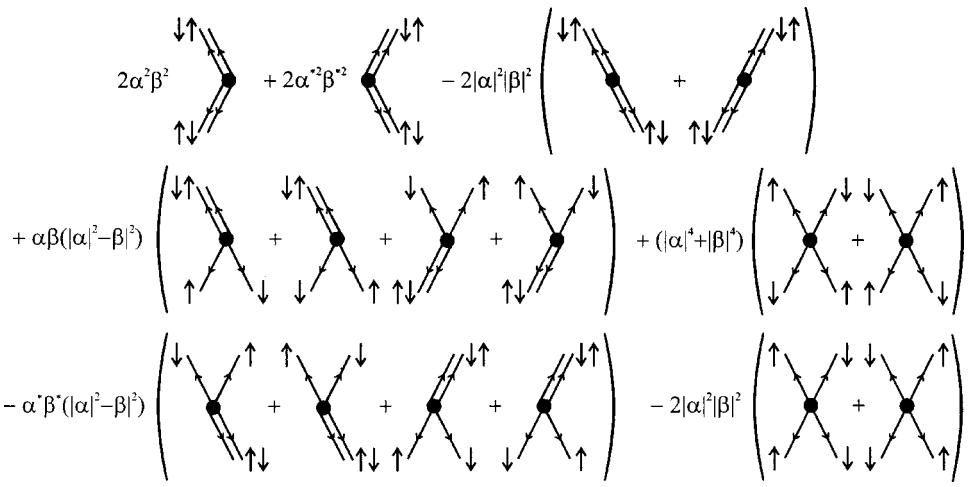


FIG. 2. Spin $|S_z|=0$ component of the total output wave function for the $(+,+)$ case, both for fermions and bosons.

We have to ensure that particles arrive at the beam splitter at the same time. The initial entanglement is between sides 1 and 2. In each pair, the particles fly apart and meet a particle from the other pair at a beam splitter. The paths on the left-hand side are labeled A and C, respectively, before and after the beam splitter. Similarly, paths on the right-hand side are labeled B and D.

The output states of this setup represent particles in paths $C1$, $D1$, $C2$, and $D2$ with a particular spin state (we note, for instance, that we can have two particles in $C1$ and none in $D1$). Now we show that, although the initial entanglement is only in the internal degrees of freedom, in the final state some of the entanglement has been transferred to the paths. We will refer to this effect as the *spin-space entanglement transfer* by local actions only.

In order to calculate what happens in the above setup, we write our initial state in the second quantization formalism,

$$\frac{1}{\sqrt{2}}(a^\dagger_{A1\uparrow}a^\dagger_{A2\downarrow}\pm a^\dagger_{A1\downarrow}a^\dagger_{A2\uparrow})\frac{1}{\sqrt{2}}(a^\dagger_{B1\uparrow}a^\dagger_{B2\downarrow}\pm a^\dagger_{B1\downarrow}a^\dagger_{B2\uparrow})|0\rangle, \tag{1}$$

where $|0\rangle$ is the vacuum state and, for instance, $a^\dagger_{A1\uparrow}$ is a creation operator describing a particle in path $A1$ and with spin up. The positive and negative signs in the above equation are necessary in order to take into account all the possible initial states (the singlet and the entangled triplet of spin). We restrict our attention to analyze one mode per particle only, but our results can be generalized to any number of modes. Due to the symmetry of the problem we only analyze two cases: when the two signs in Eq. (1) are the same, the $(+,+)$ case, and when they are different, the $(+,-)$ case. Note that initially there is no left-right correlation between spin and space. This is because there is no uncertainty in either the spin or space in the initial state, so by measuring the spatial state one cannot gain any information about the spin state (and vice versa) in addition to what we knew before the measurement.

The operation of the beam splitter is described by any unitary transformation in U(2) [13]. However, since the overall phase factor has no relevance for entanglement, we can without any loss of generality consider a transformation in SU(2):

$$U=\begin{bmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{bmatrix}, \tag{2}$$

where $|\alpha|^2+|\beta|^2=1$. Since we consider entanglement only between sides 1 and 2, the beam splitters, in fact, perform local unitary operations. Hence they cannot change the total entanglement present initially. Also, they only affect the spatial degrees of freedom and are not intrinsically dependent on spin (or polarization). Therefore, they are incapable of swapping entanglement from spin (polarization) to space by performing a controlled-NOT operation in the usual fashion [12]. Although the transformation law will be the same for fermions and bosons, they obey different statistics, which is why

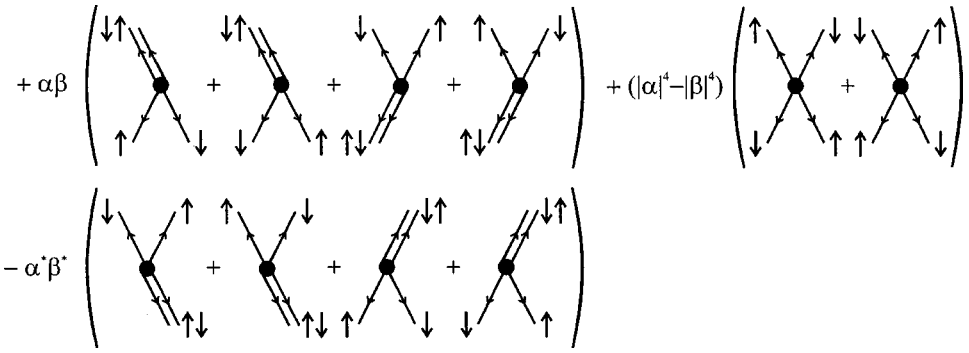


FIG. 3. Spin $|S_z|=0$ component of the total output wave function for the $(+,-)$ case, both for fermions and bosons.

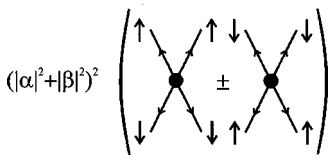


FIG. 4. Spin $|S_z|=1$ component of the total output wave function for the $(+,\pm)$ cases, for fermions.

there will be an observable difference in their behavior in our experiment. For fermions we have the following anticommutation relation:

$$[a_i^\dagger, a_j^\dagger]_+ = 0, \tag{3}$$

while for bosons we have the commutation relation:

$$[a_i^\dagger, a_j^\dagger]_- = 0, \tag{4}$$

where i and j are sets of labels. Figures 2–5 present diagramatically the output states for both fermions and bosons. For instance, the first diagram in Fig. 4 represents the following term:

$$(|\alpha|^2+|\beta|^2)^2(a_{C1\uparrow}^\dagger a_{D1\uparrow}^\dagger a_{C2\downarrow}^\dagger a_{D2\downarrow}^\dagger)|0\rangle. \tag{5}$$

Note that for each output pair, i.e., both on sides 1 and 2, the total spin (or polarization) S can take the values 0 or 1. If we consider, without any loss of generality, that the spin is aligned with the z axis, then $|S_z|$ – the absolute value of the projection of **S** along z – can also only take the values 0 or 1. We can then divide the total output wave function into these two components, where the spins of the particles in each output pair are, respectively, antialigned or aligned along z.

(1) $|S_z|=0$ component. There is no difference between fermions and bosons (bearing in mind that the corresponding operators obey different commutation relations). However, there is a difference between the $(+,+)$ case, where we have all possible output terms (see Fig. 2), and the $(+,-)$ case, where some terms never appear (see Fig. 3).

(2) $|S_z|=1$ component. There is a difference between the output states for fermions (see Fig. 4) and bosons (see Fig. 5). For both types of particles, the $(+,-)$ case will only introduce a phase difference in some terms.

As a consequence of applying only local unitary operations, the total output wave function should have also two ebits of entanglement. For clarity, let us consider for the rest of the paper the particular case of 50/50 beam splitters ($\alpha=1/\sqrt{2}, \beta=-i/\sqrt{2}$). To illustrate the spin-space entanglement transfer effect, we look at the $(+,+)$ case for fermions (Figs. 2 and 4). Here, it is clear that the $|S_z|=1$ terms give one ebit of entanglement, solely in the internal degrees of freedom, as the path states are identical. The $|S_z|=0$ case gives the other ebit of entanglement, but this time involving both the internal and external degrees of freedom. *Thus we have spin-space entanglement transfer, without any controlled operation between spin and space.*

We now show how we can extract space-only entanglement from the total wave function by doing particular measurements on the internal degrees of freedom without revealing any knowledge about the external ones (this is perfectly allowed by quantum mechanics and can be accomplished by passing the particles on each side through a cavity that extends over both the left and the right paths). For example, we can measure the total spin **S** on both sides (1 and 2) along the x axis and then select the $S_x=0$ results. For fermions, the entire wave function is then projected onto

$$\frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}(|L\rangle_1+|R\rangle_1)\frac{1}{\sqrt{2}}(|L\rangle_2+|R\rangle_2)-\frac{1}{\sqrt{2}}|A\rangle_1\frac{1}{\sqrt{2}}|A\rangle_2,\right. \tag{6}$$

where $|L\rangle_{1,2}$ means left bunching of the particles, respectively, for sides 1 and 2, $|R\rangle_{1,2}$ right bunching and $|A\rangle_{1,2}$ represents antibunching (unormalized state). The bosonic counterpart of the above state is

$$\frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}(|L\rangle_1+|R\rangle_1)\frac{1}{\sqrt{2}}(|L\rangle_2+|R\rangle_2)+\frac{1}{\sqrt{2}}|A\rangle_1\frac{1}{\sqrt{2}}|A\rangle_2.\right. \tag{7}$$

Both these states have 1 ebit of entanglement in space and the same outcome probability of 1/2. Note that since these

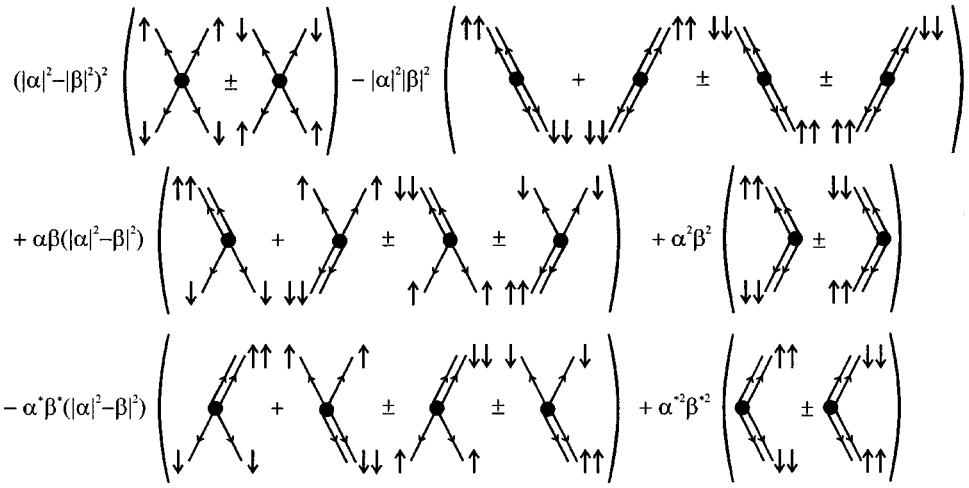


FIG. 5. Spin $|S_z|=1$ component of the total output wave function for the $(+,\pm)$ cases, for bosons.

two states are orthogonal they can be perfectly discriminated, offering an operational way of distinguishing fermions and bosons.

If on the other hand, we measure the spatial components of the total wave function, we will find different amounts of entanglement in the internal degrees of freedom of fermions and bosons. For instance, if we select the antibunching results, we will obtain the following state for fermions:

$$\frac{1}{\sqrt{3}}\left[\frac{1}{\sqrt{2}}(a^{\dagger}_{C1\uparrow}a^{\dagger}_{D1\downarrow}+a^{\dagger}_{C1\downarrow}a^{\dagger}_{D1\uparrow})\frac{1}{\sqrt{2}}(a^{\dagger}_{C2\uparrow}a^{\dagger}_{D2\downarrow}+a^{\dagger}_{C2\downarrow}a^{\dagger}_{D2\uparrow})\right.$$
$$\left.-(a^{\dagger}_{C1\uparrow}a^{\dagger}_{D1\uparrow}a^{\dagger}_{C2\downarrow}a^{\dagger}_{D2\downarrow})-(a^{\dagger}_{C1\downarrow}a^{\dagger}_{D1\downarrow}a^{\dagger}_{C2\uparrow}a^{\dagger}_{D2\uparrow})\right]|0\rangle, \quad (8)$$

with an outcome probability of 2/3 and $\log_2 3$ units of entanglement, whereas for bosons we will get

$$\frac{1}{\sqrt{2}}(a^{\dagger}_{C1\uparrow}a^{\dagger}_{D1\downarrow}-a^{\dagger}_{C1\downarrow}a^{\dagger}_{D1\uparrow})\frac{1}{\sqrt{2}}(a^{\dagger}_{C2\uparrow}a^{\dagger}_{D2\downarrow}-a^{\dagger}_{C2\downarrow}a^{\dagger}_{D2\uparrow})|0\rangle, \quad (9)$$

with probability 1/3 and 0 units of entanglement.

If we select the bunching results, for fermions we will obtain the state

$$\frac{1}{\sqrt{2}}(a^{\dagger}_{C1\uparrow}a^{\dagger}_{C1\downarrow}+a^{\dagger}_{D1\uparrow}a^{\dagger}_{D1\downarrow})\frac{1}{\sqrt{2}}(a^{\dagger}_{C2\uparrow}a^{\dagger}_{C2\downarrow}+a^{\dagger}_{D2\uparrow}a^{\dagger}_{D2\downarrow})|0\rangle, \quad (10)$$

with an outcome probability of 1/3 and 0 units of entanglement, and for bosons we will get

$$\frac{1}{\sqrt{3}}\left[\frac{1}{\sqrt{2}}(a^{\dagger}_{C1\uparrow}a^{\dagger}_{C1\downarrow}+a^{\dagger}_{D1\uparrow}a^{\dagger}_{D1\downarrow})\frac{1}{\sqrt{2}}(a^{\dagger}_{C2\uparrow}a^{\dagger}_{C2\downarrow}+a^{\dagger}_{D2\uparrow}a^{\dagger}_{D2\downarrow})\right. \quad (11)$$
$$\left.+\frac{1}{2}(a^{\dagger}_{C1\uparrow}a^{\dagger}_{C1\uparrow}+a^{\dagger}_{D1\uparrow}a^{\dagger}_{D1\uparrow})(a^{\dagger}_{C2\downarrow}a^{\dagger}_{C2\downarrow}+a^{\dagger}_{D2\downarrow}a^{\dagger}_{D2\downarrow})\right.$$
$$\left.+\frac{1}{2}(a^{\dagger}_{C1\downarrow}a^{\dagger}_{C1\downarrow}+a^{\dagger}_{D1\downarrow}a^{\dagger}_{D1\downarrow})(a^{\dagger}_{C2\uparrow}a^{\dagger}_{C2\uparrow}+a^{\dagger}_{D2\uparrow}a^{\dagger}_{D2\uparrow})\right]|0\rangle,$$

with probability 2/3 and $\log_2 3$ units of entanglement. We observe that for a given path selection one type of particles exhibits some entanglement in the internal degrees of freedom, whereas the other exhibits none. In other words, under the *same* situation, fermions and bosons show a difference in their information processing behavior. Moreover, measuring this degree of entanglement in the internal degrees of freedom could thus also be an operational way of distinguishing between fermions and bosons.

In this paper we have shown that it is possible to transfer entanglement from the internal to the spatial degrees of freedom through local actions using only the effects of particle indistinguishability and quantum statistics, without any interaction between the spin and the path. Moreover, subensembles selected by local measurements of the path will, in general, have different amounts of entanglement in the internal degrees of freedom depending on the statistics (either fermionic or bosonic) of the particles involved. This establishes a connection between two fundamental notions of quantum physics: entanglement and particle statistics. We intend to present a more detailed and systematic analysis of this setup in a subsequent longer work.

Our analysis suggests further investigation of the consequences and applications of particle statistics in quantum information processing. For example, in some protocols using spin-space entanglement the statistical effects make it unnecessary to have controlled operations, such as using polarization-dependent beam splitters [14]. Other types of statistics (e.g., anyons) can similarly be addressed within our framework. Recent experiments, such as, Refs. [15,16] suggest that it would be possible to test our results in the near future.

Y.O. acknowledges support from Fundação para a Ciência e a Tecnologia from Portugal. N.P. thanks Elsag S.p.A. for finantial support. V.V. acknowledges support from the Hewlett-Packard company, EPSRC, and the European Union project EQUIP.

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[2] E. Schrödinger, Naturwissenschaften **23**, 807 (1935); **23**, 823 (1935); **23**, 844 (1935).
[3] J. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987).
[4] See for example, A. Steane, Rep. Prog. Phys. **61**, 117 (1998).
[5] P. W. Shor, in *Proceedings of the* 35*th Annual Symposium on Foundations of Computer Science*, edited by S. Goldwasser (IEEE Computer Society Press, Santa Fe, NM, 1994), pp. 124–134; for a review of this algorithm see A. Ekert and R. Jozsa, Rev. Mod. Phys. **68**, 733 (1996).
[6] A.K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[7] S. Bose and D. Home, Phys. Rev. Lett. **88**, 050401 (2002).
[8] D.S. Abrams and S. Lloyd, Phys. Rev. Lett. **79**, 2586 (1997).
[9] J. Schliemann, J.I. Cirac, M. Kuś, M. Lewenstein, and D. Loss, Phys. Rev. A **64**, 022303 (2001); J. Schliemann, D. Loss, and A.H. MacDonald, Phys. Rev. B **63**, 085311 (2001).
[10] A.T. Costa, Jr. and S. Bose, Phys. Rev. Lett. **87**, 277901 (2001).
[11] F. Herbut and M. Vujičić, J. Phys. A **20**, 5555 (1987); F. Herbut, Am. J. Phys. **69**, 207 (2001).
[12] M. Zukowski, A. Zeilinger, M.A. Horne, and A.K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993); S. Bose, V. Vedral, and P.L. Knight, Phys. Rev. A **57**, 822 (1998).
[13] R. Loudon, Phys. Rev. A **58**, 4904 (1998).
[14] N. Paunković, Y. Omar, S. Bose, and V. Vedral, Phys. Rev. Lett. **88**, 187903 (2002).
[15] T.J. Herzog, P.G. Kwiat, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **75**, 3034 (1995).
[16] R.C. Liu, B. Odom, Y. Yamamoto, and S. Tarucha, Nature (London) **391**, 263 (1998).

Information geometry of quantum critical submanifolds: Relevant, marginal, and irrelevant operators

Bruno Mera [*]
Advanced Institute for Materials Research (WPI-AIMR), Tohoku University, Sendai 980-8577, Japan

Nikola Paunković [†]
Instituto de Telecomunicações and Departmento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal

Syed Tahir Amin [‡]
Instituto de Telecomunicações, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal;
Departamento de Física, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal;
CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal;
Department of Physics and I3N, University of Aveiro, 3810-193 Aveiro, Portugal;
Center of Physics of University of Minho and University of Porto, P-4169-007 Oporto, Portugal;
and Department of Physics, Kohsar University Murree, Murree, Pakistan

Vítor R. Vieira [§]
Departamento de Física, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal and CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal

(Received 4 January 2022; revised 3 August 2022; accepted 20 September 2022; published 3 October 2022)

We analyze the thermodynamical limit of the quantum metric along critical submanifolds of theory space. Building upon various results previously known in the literature, we relate its singular behavior to normal directions, which are naturally associated with relevant operators in the renormalization group sense. We formulate these results in the language of information theory and differential geometry. We exemplify our theory through the paradigmatic examples of the XY and Haldane models, where the normal directions to the critical submanifolds are seen to be precisely those along which the metric has singular behavior, while for the tangent ones it vanishes—these directions lie in the kernel of the metric.

DOI: 10.1103/PhysRevB.106.155101

I. INTRODUCTION

Classical information geometry is the study of the geometry of statistical manifolds, and it finds applications in statistics, information theory, and statistical mechanics [1–3]. Recently, information geometry has seen a lot of application in the study of quantum systems and their phase transitions by analyzing different Riemannian metrics over spaces of density operators, a generalization of classical probability distributions [4–8] (for more details on quantum information geometry and applications, see, for example, Refs. [9–17]). Of particular interest are the exotic topological phases of matter that go beyond the standard Landau-Ginzburg classification paradigm and for which there is no well-defined local order parameter, used to infer phase transitions. The quantum metric [18]—a Riemannian metric over the space of pure quantum states based on state distinguishability—was shown to be a quantity able to probe zero-temperature quantum phase transitions among these phases as one varies some parameter in the system [19–21]. In the context of Bloch bands, the quantum metric in momentum space, in the particular setting of flat bands, has received a lot of attention recently, as it gives geometrical contributions to the characterization of a range of different phenomena, such as exotic superconductivity and superfluidity [22–25], the stability of fractional Chern insulating phases [26–31], and light-matter coupling [32]. The quantum metric is also central in determining maximally localized Wannier functions [33,34], and it can be used as a practical indicator for exotic momentum-space monopoles [35,36]. The integral of the momentum-space quantum metric is a measure of electron localization, and it can be extracted through spectroscopy measurements [37]. In two spatial dimensions, one can define an associated complex structure, which is a measure of anisotropy in localization [38].

More recently, relations between the quantum metric and the Berry curvature, which determines topological invariants of the system and gives rise to Berry-phase effects, have been established and understood based on the Kähler structure of

[*]bruno.mera.c5@tohoku.ac.jp
[†]npaunkov@math.tecnico.ulisboa.pt
[‡]tahiramin811@gmail.com
[§]vitor.rocha.vieira@tecnico.ulisboa.pt

 ©2022 American Physical Society

the space of quantum states [39,40]. These relations have come to play an important role in recent studies concerning the so-called ideal Chern bands, which are presumed to be ideal candidates for hosting fractional Chern insulating phases, and the associated band structure engineering and transport [31,41–46]. More general relations between the quantum metric and topology of quantum states were derived in the context of Dirac Hamiltonians [47].

The quantum metric over the space of parameters describing the system, being the "infinitesimal" distance between two ground states, allows us to probe phase transitions. As long as the quantum metric is regular, the ground state of the system does not change substantially, keeping the system in the same phase. On the other hand, if the quantum metric becomes singular, it means that the ground state has undergone a dramatic change and thus the system has experienced a phase transition. However, this singular behavior has its subtleties regarding the scaling behavior when performing the thermodynamical infinite-volume limit. In the following, we will show how the quantum metric behaves along a submanifold of parameter space composed of critical points—described, in the long-wavelength and infinite-volume limit, by conformal field theories. We will show that there exist two types of directions of the parameter change: those which take the system away from, and those which move it within the critical submanifold. By performing a scaling analysis, we show that the behaviors of these directions can be interpreted in light of the theory of the renormalization group (RG); namely, the former directions correspond to relevant operators in the quantum theory, while the latter correspond to irrelevant and marginal operators in the quantum theory. While most of these results are known in the literature, they appear to be scattered throughout a number of papers, done by researchers working in different fields of physics who may not necessarily be aware of each other's results. Here, we present a unified formulation of those results by describing them through the language of information theory and differential geometry. We also illustrate our results on the examples of the XY and the Haldane models. Finally, we present conclusions and future lines of research.

II. QUANTUM METRIC ALONG A CRITICAL SUBMANIFOLD

Suppose we are given a many-body Hamiltonian depending smoothly on some parameters, collectively denoted by x, living in a smooth manifold $\mathcal{T}$ of dimension n. We refer to $\mathcal{T}$ as *theory space*. These parameters will typically correspond to couplings and external fields determining the system's Hamiltonian $H(x)$, with $x \in \mathcal{T}$. We will be interested in the ground state, i.e., the zero-temperature properties of the system. In particular, we consider the pullback of the *Fubini-Study metric* (which is the natural unitarily invariant metric in the space of pure quantum states), also known as the fidelity susceptibility or the information metric, with respect to the family of ground states $\rho(x) = |\psi(x)\rangle\langle\psi(x)|$ of $H(x)$, with $x \in \mathcal{T}$. Note that there may be parameters for which the ground state is not unique, in which case we will adopt a regularization procedure where we consider the pullback of the *Bures metric*—a Riemannian metric on the space of density operators—under the map $x \mapsto e^{-H(x)/T}/\mathrm{Tr}(e^{-H(x)/T})$ and eventually take the $T \to 0$ limit. Note that, in order to be able to infer the phase transitions, one must first take the $T \to 0$ limit before the thermodynamic limit, since the two do not commute, as discussed in Ref. [21]. Only then is the regularization procedure meaningful, giving consistent results and treating nondegenerate and degenerate cases equally. Additionally, the Boltzmann-Gibbs distribution will always treat degenerate ground states equally, because they have the same energy, which is reasonable to assume given no additional information. The Bures metric has two independent contributions coming from the "classical" and "quantum" parts of the density matrix $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$, corresponding to the variations on the probabilities and the eigenstates of ρ, respectively; see Ref. [6]:

$$ds^2 = \frac{1}{4}\sum_j \frac{dp_j^2}{p_j} + \frac{1}{2}\sum_{j\neq i} |\langle\psi_j|d|\psi_i\rangle|^2 \frac{(p_j - p_i)^2}{p_i + p_j}. \quad (1)$$

The first term is the classical Fisher metric corresponding to the statistical model determined by the probability distribution (p_j). The second term, measuring the variations in the eigenstates of ρ, has the nice property that for pure states $\rho = |\psi\rangle\langle\psi|$ it reduces to the Fubini-Study metric

$$g_{\mathrm{FS}} = \langle d\psi|(1 - |\psi\rangle\langle\psi|)|d\psi\rangle. \quad (2)$$

Under the map $x \mapsto |\psi_i(x)\rangle\langle\psi_i(x)|$, where $|\psi_i(x)\rangle$ is an eigenstate of $H(x)$ (assuming such a state is globally defined, up to a phase), the pullback of the Fubini-Study metric g_{FS} describes a metric in the space of theories $\mathcal{T}$—the *quantum metric*, which has the nice formula in terms of the remaining eigenstates

$$g = \sum_{j\neq i} \langle d\psi_i(x)|\psi_j(x)\rangle\langle\psi_j(x)|d\psi_i(x)\rangle. \quad (3)$$

We will now take $|\psi_i(x)\rangle = |\psi(x)\rangle$ to be the ground state of $H(x)$. It is convenient to interpret the quantum metric g in terms of operators acting on the Hilbert space. At a given point $x \in \mathcal{T}$, the metric assigns a bilinear non-negative pairing $g(u, v)$ of tangent vectors $u, v \in T_x\mathcal{T}$. Physically, an element $v \in T_x\mathcal{T}$ generates a first-order variation in the parameters of the theory at $x \in \mathcal{T}$. As such, it has a corresponding operator, $\mathcal{O}_v$, acting on the Hilbert space, which, using local coordinates x^i, $i = 1, \ldots, n$, where $v = v^i \frac{\partial}{\partial x^i}$, is given by

$$\mathcal{O}_v = v^i \frac{\partial H}{\partial x^i}, \quad (4)$$

where we assume the Einstein summation convention. It is clear that the operator $\mathcal{O}_v$ is associated with perturbation theory in the sense that $H(x + \varepsilon v) = H(x) + \varepsilon\mathcal{O}_v + \mathrm{O}(\varepsilon^2)$. At each parameter point x from our space of theories $\mathcal{T}$, once we fix local coordinates x^i, $i = 1, \ldots, n$, we have a basis for an associated linear space of operators given by $\mathcal{O}_i(x) = \mathcal{O}_{\partial/\partial x^i}$, $i = 1, \ldots, n$. The pullback of the Fubini-Study metric under the map $x \mapsto |\psi(x)\rangle\langle\psi(x)|$ describes the quantum metric in the space of theories $\mathcal{T}$. Note, however, that it may happen that the metric is degenerate, meaning it is not necessarily invertible at every point x where it is defined—simply because the Jacobian of the transformation can be singular. Alternatively, we can view the degeneracy of the metric in light of the

tangent-vector-to-operator correspondence $v \mapsto \mathcal{O}_v$. Assuming the eigenstates are nondegenerate in energy, let $|\psi_i(x)\rangle$ be the eigenstates for $H(x)$ with corresponding energies $E_i(x)$. The eigenstates of $H(x+\varepsilon v)$ can be obtained from first-order perturbation theory around $H(x)$ as follows:

$$|\psi_i(x+\varepsilon v)\rangle = |\psi_i(x)\rangle + \varepsilon \sum_{j\neq i} \frac{\langle\psi_j(x)|\mathcal{O}_v|\psi_i(x)\rangle}{E_i(x)-E_j(x)}|\psi_j(x)\rangle, \tag{5}$$

where we assume a parallel transport gauge $\langle\psi_i(x)|v^k\frac{\partial}{\partial x^k}|\psi_i(x)\rangle = 0$. The order ε term can vanish if $v \in T_x\mathcal{T}$ is in the kernel of the assignment $v \mapsto \mathcal{O}_v$. More generally, it can vanish if and only if $\mathcal{O}_v$ preserves the eigenspace $L_x^{(i)} = \mathrm{span}_{\mathbb{C}}\{|\psi_i(x)\rangle\}$, the fiber at x of the ith eigenbundle of H. Observe that from Eq. (3), it follows that for any $u \in T_x\mathcal{T}$,

$$g(v,u) = \frac{1}{2}\sum_{j\neq i}\frac{\langle\psi_i|\mathcal{O}_v|\psi_j\rangle\langle\psi_j|\mathcal{O}_u|\psi_i\rangle}{(E_i(x)-E_j(x))^2} + (v \leftrightarrow u), \tag{6}$$

and we see that $g(v,\cdot) = 0$, i.e., $g(v,u) = 0$ for all $u \in T_x\mathcal{T}$ if and only if $\mathcal{O}_v$ preserves the eigenspace $L_x^{(i)}$. In short, for the ground state $\rho(x) = |\psi(x)\rangle\langle\psi(x)|$, the above statement can be expressed as

$$[\mathcal{O}_v, \rho(x)] = 0 \iff g(v,\cdot) = 0. \tag{7}$$

From a mathematical perspective, we can understand the above statement as follows. There are two points to be considered here. One is that the smooth map $H : \mathcal{T} \ni x \mapsto H(x)$, where $H(x)$ belongs to the (real) vector space of Hermitian operators, has a differential dH, and what we call $\mathcal{O}_i$ is the pushforward $dH(\frac{\partial}{\partial x^i})$, which makes sense on a local chart where the local coordinates x^i's are defined. The map dH can be seen as a vector bundle map between the tangent bundle $T\mathcal{T}$ and the pullback under H of the tangent bundle of the (real) vector space of Hermitian operators. The map dH is not necessarily injective, so one must be careful in identifying the tangent space $T_x\mathcal{T}$ with the image of dH at a point.

The second point is that the smooth assignment $P : x \mapsto P(x) = |\psi_i(x)\rangle\langle\psi_i(x)|$ (provided the ith eigenline bundle is well defined), where we identify the target space as the projectivization of the Hilbert space, also has a differential which is given explicitly by

$$dP = QdPP + PdPQ, \tag{8}$$

where $Q = I - P = \sum_{j\neq i}|\psi_j\rangle\langle\psi_j|$ is the orthogonal complement projector (here, the $|\psi_j\rangle$'s are locally defined orthonormal eigenvectors of H). Above we used the fact that $P^2 = P \Rightarrow PdPP = 0$. Observe that

$$QdPP = \sum_{j\neq i}\langle\psi_j|d|\psi_i\rangle|\psi_j\rangle\langle\psi_i| = (PdPQ)^\dagger \tag{9}$$

and that (under the assumption $E_i \neq E_j$, for $i \neq j$)

$$H|\psi_i\rangle = E_i|\psi_i\rangle \Rightarrow \frac{\langle\psi_j|dH|\psi_i\rangle}{E_i-E_j} = \langle\psi_j|d|\psi_i\rangle. \tag{10}$$

It follows that

$$QdPP = \sum_{j\neq i}\langle\psi_j|d|\psi_i\rangle|\psi_j\rangle\langle\psi_i| = \sum_{j\neq i}\frac{\langle\psi_j|dH|\psi_i\rangle}{E_i-E_j}|\psi_j\rangle\langle\psi_i|. \tag{11}$$

The differential of P, the injectivity of which determines the nondegeneracy of the quantum metric (since the Fubini-Study metric is a Riemannian and hence nondegenerate metric on the projectivization of the Hilbert space), is controlled by the differential of H. Observe that for a tangent vector v in $\mathcal{T}$ we have $(QdPP)(v) = 0 \iff dP(v) = 0$ (because $QdPP$ and $PdPQ$ are orthogonal) and also $dP(v) \iff P^*g_{\mathrm{FS}}(v,\cdot) = 0$. Now the only way $(QdPP)(v)$ vanishes is if $dH(v) = \mathcal{O}_v$ preserves the state $|\psi_i\rangle$.

We can relate the above discussion to the so-called *symmetric logarithmic derivative equation* as follows. The symmetric logarithmic derivative of a family of density operators is determined by an operator valued $1-$form G solving the symmetric logarithmic derivative equation

$$d\rho = G\rho + \rho G. \tag{12}$$

In the particular case of $\rho = P = |\psi_i\rangle\langle\psi_i|$, we can see that

$$G = QdPP + PdPQ \tag{13}$$

solves Eq. (12), and hence the vanishing of $(QdPP)(v)$, for a tangent vector v, is equivalent to the vanishing of $G(v)$, known as the symmetric logarithmic derivative operator of $\rho = P$ with respect to v.

Suppose now we have a set $C \subset \mathcal{T}$ of *quantum critical points* where the theory is gapless and our geometric quantity of interest g is presumably singular in the thermodynamical limit. Note that the notion of singularity here is different from the familiar notion of singularity of a smooth vector field in the context of differential geometry. Here the singularity is associated with the quantum metric. More precisely, we say that the quantum metric has a singularity at a point $x \in \mathcal{T}$ if, by taking local coordinates centered at x and hence obtaining a matrix representation of it (by considering the natural coordinate tangent vectors), some of the associated matrix elements are not defined at that point. Note that this notion is really coordinate independent: If this happens in a coordinate system, it will happen in any coordinate system. The thermodynamical limit assumes, in particular, that the volume of the system L^d, where L is the linear length of the system and d is the spatial dimension, goes to infinity. For considerations involving the thermodynamical limit, it will be convenient to consider the rescaled quantum metric g/L^d. We also make the assumption that the quantum critical theories, as described by points $x \in C \subset \mathcal{T}$, are conformally invariant so that they are described, in the continuum limit and at large distances, by a conformal field theory (CFT_d). For simplicity, we assume that C is a submanifold of $\mathcal{T}$, and we refer to it as a *critical submanifold*. Note that in the model examples below, there will be regions of theory space which are critical in the same sense, i.e., the Hamiltonian is gapless in those regions; however, these regions will fail to be submanifolds. We will refer to these regions as *critical regions*. Of particular interest to us will be the tangent bundle of C in $\mathcal{T}$, i.e., the vector bundle over C that at each point $x \in C$ associates the

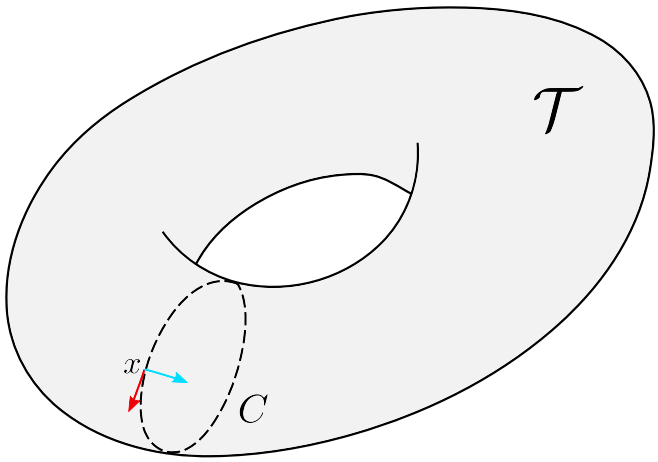


FIG. 1. An illustration of the space of theories $\mathcal{T}$ (here, a compact surface of genus 1), together with a quantum critical submanifold $C \subset \mathcal{T}$ (here, a circle). The red arrow is meant to illustrate a tangent vector to C at $x \in C$, which maps to a marginal or irrelevant operator. The blue arrow illustrates a normal vector to C at $x \in C$, mapping to a relevant operator.

tangent space $T_xC \subset T_x\mathcal{T}$. The point that we want to make, which will be clear from the discussion below, is that the rescaled metric g/L^d is actually, in the thermodynamical limit, only divergent along directions which are complementary to TC, which we refer to as directions of *increased ground-state distinguishability* and are associated with *relevant operators* in the framework of the RG. To be precise, the thermodynamical limit of g/L^d, which we denote by g_*, vanishes exactly along tangent vectors to C, i.e., $g_*(v, \cdot) = 0$ for any vector $v \in T_xC \subset T_x\mathcal{T}$. As a consequence, the directions where g_* is nonvanishing are normal to TC—i.e., each can be seen, effectively, as orthogonal to TC with respect to some Riemannian metric in $\mathcal{T}$. The rigorous statement is that, for each $x \in C$, there is an *exact sequence of vector spaces*

$$0 \longrightarrow T_xC \longrightarrow T_x\mathcal{T} \longrightarrow N_xC \longrightarrow 0, \tag{14}$$

defining the vector space of normal directions at x, $N_xC = T_x\mathcal{T}/T_xC$, giving rise, globally, to the normal bundle $NC = T\mathcal{T}|_C/TC \to C$; see Fig. 1 (by an exact sequence of vector spaces, we mean a sequence of vector spaces and linear maps where the kernel of the next one is equal to the image of the current one). Observe that for $x \in C$ and any $u \in T_x\mathcal{T}$, if $g_*(u, u) \neq 0$, we have that $g_*(u + v, u) = g_*(u, u)$ for any $v \in T_xC$, since v is annihilated by g_*. The vector u defines a nontrivial element $[u]$ of $N_xC = T_x\mathcal{T}/T_xC$, and $g_*(u, u)$ only depends on $[u]$. A choice of a complement of T_xC gives a particular unique representation of $[u]$ as an element of $T_x\mathcal{T}$. In practice, one chooses an auxiliary Riemannian metric h defined over $\mathcal{T}$ (actually, it is enough to have it defined over $T\mathcal{T}|_C$) and identifies $N_xC \cong (T_xC)^{\perp_h}$ using the notion of an orthogonal complement provided by h. Importantly, the isomorphism class of this vector bundle is independent of this choice. Finally, we remark that, in light of Eq. (7), we see that the operators associated with T_xC commute with the ground state at x modulo subextensive terms, i.e., terms that scale as L^r, with $r < d$.

For the convenience of the reader, before moving on to the relation with RG, we list our assumptions.

(i) The theory space or parameter space $\mathcal{T}$ is assumed to be a smooth (C^∞) manifold.

(ii) A critical submanifold C is a submanifold $C \subset \mathcal{T}$ where $H(x)$ is, in the thermodynamical limit, gapless for every $x \in C$. If C is not a submanifold (as happens in the examples considered) but $H(x)$ is, in the thermodynamical limit, gapless for every $x \in C$, we refer to it as a critical region.

(iii) Points in C are assumed to be described by conformal field theories in the continuum limit, at low energies.

III. RELATION TO RG

We consider a critical point $x_0 \in C \subset \mathcal{T}$ and assume that the $\mathcal{O}_i$'s are local operators, meaning that they can be written as

$$\mathcal{O}_i = \sum_{\mathbf{r}} \mathcal{O}_i(\mathbf{r}), \tag{15}$$

for $\mathcal{O}_i(\mathbf{r})$ having support in some bounded neighborhood of $\mathbf{r}$. The sum over lattice sites $\mathbf{r}$ makes sense before taking the continuum limit. It is convenient to introduce the imaginary-time-evolved operators, namely,

$$\mathcal{O}_i(\tau, \mathbf{r}) = e^{\tau H(x_0)} \mathcal{O}_i(\mathbf{r}) e^{-\tau H(x_0)}, \tag{16}$$

where τ denotes the imaginary time. At the critical point $x_0 \in C$, we may assume that the $\mathcal{O}_i$'s have well-defined scaling dimensions Δ_i such that

$$\mathcal{O}_i(\tau, \mathbf{r}) \longrightarrow \mathcal{O}_i'(\tau, \mathbf{r}) = \mathcal{O}_i(\zeta^z \tau, \zeta \mathbf{r}) \zeta^{\Delta_i}, \tag{17}$$

under a transformation $\mathbf{r} \to \zeta \mathbf{r}$ and $\tau \to \zeta^z \tau$, where z is the dynamical critical exponent. If this is not the case, assuming $\dim \mathcal{T}$ is big enough, we may use a linear coordinate change, so that the above equation holds. These scaling dimensions arise from the linearization of the renormalization group flow near the critical point $x_0 \in C$; see Ref. [48] for details.

Due to the expression

$$g_{ij}(x) = \int_0^\infty d\tau\, \tau e^{-\varepsilon\tau} \times \left[\frac{1}{2} \langle \{ \mathcal{O}_i(\tau), \mathcal{O}_j(0) \} \rangle - \langle \mathcal{O}_i(\tau) \rangle \langle \mathcal{O}_j(0) \rangle \right], \tag{18}$$

with $\varepsilon \to 0^+$ and $\{\cdot, \cdot\}$ being the anticommutator, one discovers that $g_{ij}(x_0)$ has the finite-size scaling law originally found by Campos Venuti, Zanardi, and co-workers [5,7] and later found in the work of Miyaji *et al.* [49], in the context of gauge-gravity duality,

$$g_{ij}(x_0) \sim L^{2d+2z-\Delta_i-\Delta_j}, \tag{19}$$

where L is the linear size of the system and d is its spatial dimension. In particular, we have the following classification of operators according to how their couplings are renormalized under an RG transformation, as we flow onto the infrared energy scales: (i) $\mathcal{O}_i$ is *relevant* if $\Delta_i < d + z$, (ii) $\mathcal{O}_i$ is *marginal* if $\Delta_i = d + z$, and (iii) $\mathcal{O}_i$ is *irrelevant* if $\Delta_i > d + z$.

We then see that for tangent vectors whose associated operators are *relevant* operators, the metric should blow up in the $L \to \infty$ thermodynamical limit. For tangent vectors

associated with *irrelevant* operators the metric should vanish in the thermodynamical limit. Finally, for tangent vectors associated with *marginal* operators, the metric should scale as L^0, meaning it should be finite, with possible logarithmic corrections. Accordingly, if we are in a point x taken from a small neighborhood of $x_0 \in C$, the relevant operators move us outside C, the irrelevant operators move us back to x_0, while the marginal operators move us within C, since the corresponding theory is also *scale invariant*. Thus we expect that the tangent space $T_{x_0}C$ can be identified with the subspace of all irrelevant and marginal operators in the image of $T_{x_0}\mathcal{T} \ni v \mapsto \mathcal{O}_v$. The marginal operators in the image of $T_{x_0}\mathcal{T} \ni v \mapsto \mathcal{O}_v$ then form a complementary subspace $N_{x_0}C$, such that $T_{x_0}\mathcal{T} = T_{x_0}C \oplus N_{x_0}C$.

Consider the rescaled metric $g_{ij}(x_0)/L^d$. The scaling behavior of its diagonal components is given by

$$\frac{g_{ii}}{L^d} \sim \frac{1}{L^d} L^{2z+2d-2\Delta_i}, \tag{20}$$

and hence it vanishes for $2z + d - 2\Delta_i < 0$ and is finite or blows up for $2z + d - 2\Delta_i \geqslant 0$ in the thermodynamic limit. In particular, for irrelevant and marginal operators, it goes to zero. Surprisingly, it also vanishes for those relevant operators that are not "sufficiently relevant," namely, those that satisfy

$$z + d > \Delta_i > z + \frac{d}{2}. \tag{21}$$

Its thermodynamical limit g_* vanishes along $T_{x_0}C$, so we can identify $T_{x_0}C$ as a subspace of $T_{x_0}\mathcal{T}$ where the rescaled metric has a nontrivial kernel. In the following, we consider two particular examples for which the tangent bundle to C in $\mathcal{T}$ is exactly the kernel of g_*—the XY and the Haldane models.

IV. XY MODEL

In the XY anisotropic spin-half chain with N sites on a circle in the presence of an external magnetic field, we have the family of Hamiltonians

$$H(\gamma, \lambda) = -\sum_{j=0}^{N-1} \left(\frac{1+\gamma}{2} \sigma_j^x \sigma_{j+1}^x + \frac{1-\gamma}{2} \sigma_j^y \sigma_{j+1}^y + \frac{\lambda}{2} \sigma_j^z \right), \tag{22}$$

parametrized by $(\gamma, \lambda) \in \mathbb{R}^2$, where γ is the anisotropy and λ is the magnetic field. This model is usually solved by using a Jordan-Wigner transformation,

$$c_j = e^{i\pi \sum_{k=0}^{j-1} \sigma_j^+ \sigma_j^-} \sigma_j^- \quad \text{and} \quad c_j^\dagger = \sigma_j^+ e^{-i\pi \sum_{k=0}^{j-1} \sigma_j^+ \sigma_j^-}, \tag{23}$$

which takes spin variables, σ_j's, to fermionic ones, c_j's, to get

$$H(\gamma, \lambda) = -\sum_{j=0}^{N-1} [(c_{j+1}^\dagger c_j + c_j^\dagger c_{j+1}) + \gamma (c_{j+1} c_j + c_j^\dagger c_{j+1}^\dagger)] - 2\lambda \sum_{j=0}^{N-1} \left(c_j^\dagger c_j - \frac{1}{2} \right), \tag{24}$$

where the fermions satisfy twisted boundary conditions $c_N^\dagger = c_0^\dagger (-1)^{\sum_j c_j^\dagger c_j}$ according to the parity operator.

One then considers the system as a whole with fixed periodic or antiperiodic boundary conditions for the fermions. The translation invariance of the Hamiltonian motivates us to Fourier-expand $c_j^\dagger = \frac{1}{\sqrt{N}} \sum_k e^{-ikj} c_k^\dagger$, where, due to the boundary conditions, we have $e^{ikN} = \pm 1$, which gives the two sets of allowed momenta. For the periodic boundary conditions the allowed momenta have the form

$$k = \frac{2\pi}{N} m, \quad m \in \{0, \ldots, N-1\}, \tag{25}$$

while for the antiperiodic boundary conditions we have

$$k = \frac{2\pi}{N} m + \frac{\pi}{N}, \quad m \in \{0, \ldots, N-1\}. \tag{26}$$

In either case, the corresponding two Hamiltonians assume the same form

$$H(\gamma, \lambda) = \sum_k [(-\lambda - \cos(k))(c_k^\dagger c_k - c_{-k} c_{-k}^\dagger) + i\gamma \sin(k) c_{-k} c_k - i\gamma \sin(k) c_k^\dagger c_{-k}^\dagger], \tag{27}$$

where we dropped the overall constant $-\sum_k \cos(k)$. It is convenient to introduce the Nambu spinor

$$\psi_k^\dagger = (c_k^\dagger \; c_{-k}), \tag{28}$$

in terms of which we have

$$H(\gamma, \lambda) = \frac{1}{2} \sum_k \psi_k^\dagger H(k) \psi_k, \tag{29}$$

with $H(k) = \vec{d}(k) \cdot \vec{\sigma}$ and $\vec{d}(k) = (0, 2\gamma \sin(k), -2(\lambda + \cos(k)))$. One can then diagonalize each Hamiltonian through a Bogoliubov-Valatin transformation to obtain the two spectra and corresponding eigenstates. By fixing periodic boundary conditions for the fermions, only the states with even parity are true eigenstates of the original Hamiltonian. Similarly, fixing antiperiodic boundary conditions, only odd-parity states are eigenstates of the original Hamiltonian. To find the true ground state of the system, one needs to consider the lowest-energy state according to this prescription, the parity of which may depend on the considered system size. In particular, it is known (see Ref. [50]) that there are phases with two degenerate ground states for $|\lambda| < 1$ and $\gamma \neq 0$, and even phases with infinitely many ground states in the thermodynamical limit (where $\lambda = 0$ and $\gamma = \pm 1$).

Instead of considering this procedure, because solving the XY model is not the aim of this paper, for simplicity, we will fix periodic boundary conditions for the fermions and take that as our model. The ground state is then easy to understand, and one can perform the standard analysis in the thermodynamic limit. The critical region then consists of three critical submanifolds given, respectively, by the segment defined by $\gamma = 0$ and $|\lambda| \leqslant 1$, and the two lines given by $|\lambda| = 1$ (see, for example, Ref. [4]). The segment and the lines intersect transversally. The whole set, i.e., the union of the critical submanifolds, is not a submanifold itself, as the intersection points do not have neighborhoods homeomorphic to $\mathbb{R}$.

For this model, we see that g_* restricted to the critical lines vanishes precisely on tangent vectors to the critical lines, except on the intersections $(\lambda, \gamma) = (\pm 1, 0)$, where it blows up in all directions. On the complementary subspaces to the tangent spaces it always blows up. In this case there are no operators which are not sufficiently relevant. We remark

that at the two critical points $(\gamma, \lambda) = (0, \pm 1)$, corresponding to the intersection between the segment and the two lines, the metric blows up along any direction, and hence they behave as if they were two isolated critical points (because the normal spaces in this case coincide with the whole of $T_{(0,\pm 1)}\mathcal{T}$). Note, however, that they are not isolated critical points in the strict sense, as any neighborhood of them finds other critical points. We refer the reader to the Appendix for details.

V. HALDANE MODEL

We consider the Haldane model Hamiltonian [51]

$$\mathcal{H} = t\sum_{\langle l,m\rangle} c_l^\dagger c_m + t'\sum_{\langle\langle l,m\rangle\rangle} e^{-i\nu_{lm}\phi} c_l^\dagger c_m + M\sum_l \epsilon_l c_l^\dagger c_l, \quad (30)$$

where $\epsilon_l = 1$ on site A and -1 on site B; $\nu_{lm} = \pm 1$ depending on the direction of next-nearest-neighbor hopping. This model supports quantized conductance without applying an external magnetic field. We also fix $t = 1$ and $t' = 1/3$. In this case, (ϕ, M) are the coordinates describing the parameter manifold in which different topological phases exist. We remark that, in this case, the manifold has the topology of an infinite cylinder $S^1 \times \mathbb{R}$, since we have to identify ϕ and $\phi + 2\pi$. The critical region for this model is well known, and in the (ϕ, M) plane they have the shape of a *figure eight*; see Ref. [51]. Once we identify the points $\phi \sim \phi + 2\pi$, the resulting critical region consists of two circles which intersect at two points described by $(M, \phi) = (0, 0)$ and $(\pi, 0)$, both intersecting transversally. To calculate the fidelity susceptibility, one considers the model on a two-dimensional (2D) lattice with L^2 sites and takes periodic boundary conditions. Then, one can write the Hamiltonian, in momentum space, which is described by a 2×2 Bloch Hamiltonian of the form $H(\mathbf{k}) = \vec{d}(\mathbf{k}) \cdot \vec{\sigma}$, where $\vec{d}(\mathbf{k})$ is a three-dimensional vector, $\vec{\sigma}$ is the vector of Pauli matrices, and $\mathbf{k}$ is the momentum in the first Brillouin zone BZ^2; see Ref. [51] (we are omitting a term proportional to the identity matrix which is irrelevant for the present discussion). Explicitly, the vector $\mathbf{d}(\mathbf{k}) = (d_1(\mathbf{k}), d_2(\mathbf{k}), d_3(\mathbf{k}))$ is given by

$$\begin{aligned} d_1(\mathbf{k}) &= 1 + \cos(\mathbf{k}\cdot\mathbf{a}_1) + \cos(\mathbf{k}\cdot\mathbf{a}_2), \\ d_2(\mathbf{k}) &= \sin(\mathbf{k}\cdot\mathbf{a}_1) + \sin(\mathbf{k}\cdot\mathbf{a}_2), \\ d_3(\mathbf{k}) &= M + \tfrac{2}{3}\sin(\phi)[\sin(\mathbf{k}\cdot\mathbf{a}_1) - \sin(\mathbf{k}\cdot\mathbf{a}_2) \\ &\quad - \sin(\mathbf{k}\cdot(\mathbf{a}_1 - \mathbf{a}_2))], \end{aligned} \quad (31)$$

with $\mathbf{a}_1 = (a/2)(3, \sqrt{3})$ and $\mathbf{a}_2 = (a/2)(3, -\sqrt{3})$, where a is the lattice constant that we take to be equal to 1. The quantum metric over the parameter space is then given by

$$g(M,\phi) = \frac{1}{4}\sum_{\mathbf{k}}\left(\frac{\partial\vec{n}}{\partial M}\cdot\frac{\partial\vec{n}}{\partial M}dM^2 + 2\frac{\partial\vec{n}}{\partial M}\cdot\frac{\partial\vec{n}}{\partial\phi}d\phi dM + \frac{\partial\vec{n}}{\partial\phi}\cdot\frac{\partial\vec{n}}{\partial\phi}d\phi^2\right), \quad (32)$$

where $\vec{n}(\mathbf{k}) = \vec{d}(\mathbf{k})/|\vec{d}(\mathbf{k})|$, and the sum is restricted to the allowed momenta for the corresponding finite-size system which yields a discrete torus inside the Brillouin zone—more explicitly, these momenta are represented in $\mathbb{R}^2$ by $\mathbf{k} = \sum_{i=1}^2 k_i \mathbf{e}_i$, where $k_i \in \{\frac{2\pi j}{L} : j = 0, \ldots, L-1\}$, $i = 1, 2$, and $\{\mathbf{e}_i\}_{i=1}^2$ is a basis for the reciprocal lattice. The fidelity susceptibility is then $\chi = g/L^2$. We remark that the expression for χ defines a Riemann sum for an integral and, once we take $N \to \infty$, we obtain the formula

$$\chi(M,\phi) = \frac{1}{4}\int_{\mathrm{BZ}^2}\frac{d^2\mathbf{k}}{(2\pi)^2}\left(\frac{\partial\vec{n}}{\partial M}\cdot\frac{\partial\vec{n}}{\partial M}dM^2 + 2\frac{\partial\vec{n}}{\partial M}\cdot\frac{\partial\vec{n}}{\partial\phi}d\phi dM + \frac{\partial\vec{n}}{\partial\phi}\cdot\frac{\partial\vec{n}}{\partial\phi}d\phi^2\right), \quad (33)$$

which can safely be used away from the critical points. For numerical purposes, one can use the procedure outlined in Sec. IV of Ref. [21] to compute g and χ for a finite system and a very small, but nonvanishing, temperature.

In Fig. 2 we present fidelity susceptibilities (rescaled quantum metric diagonal components) $\chi_{MM} = g_{MM}/L^2$ (the upper plots) and $\chi_{\phi\phi} = g_{\phi\phi}/L^2$ (the middle plots) calculated at temperature $T = 2 \times 10^{-5}$ and $N_x = N_y = 30$ sites. We see that both plots clearly show the critical lines, apart from parts of the second plot when the change of the parameter ϕ is precisely tangent to the critical manifold, showing that the tangent directions have decreased distinguishability. Just like the case of the XY model, the points $(M, \phi) = (0, 0)$ and $(0, \pi)$ behave as isolated critical points, and the metric blows up in all directions in the thermodynamic limit.

The off-diagonal contribution to the metric shows a very interesting behavior, which we proceed to explain. First, it is consistent with what is inferred from the diagonal contributions, as it also singles out the critical region. Secondly, it gives us additional information about the tangent spaces to the critical submanifold. Namely, we see that $\chi_{M\phi}$ changes sign along the critical region in six points, where it is necessarily zero. The reason for this is twofold. Recall that $\chi_{M\phi}$ is given by the scalar product between $\frac{\partial}{\partial M}$ and $\frac{\partial}{\partial\phi}$. At the "isolated critical points" $(M, \phi) = (0, 0)$ and $(0, \pi)$, this scalar product becomes zero, and the two vectors become orthogonal. On the other hand, at the other four points, characterized by $\phi = \pm\pi/2$, the vector $\frac{\partial}{\partial\phi}$ is precisely tangent to the critical region, which can be inferred from the plot of $\chi_{\phi\phi}$.

To further analytically confirm the behaviors exhibited in Fig. 2, one would need to perform careful expansions, as

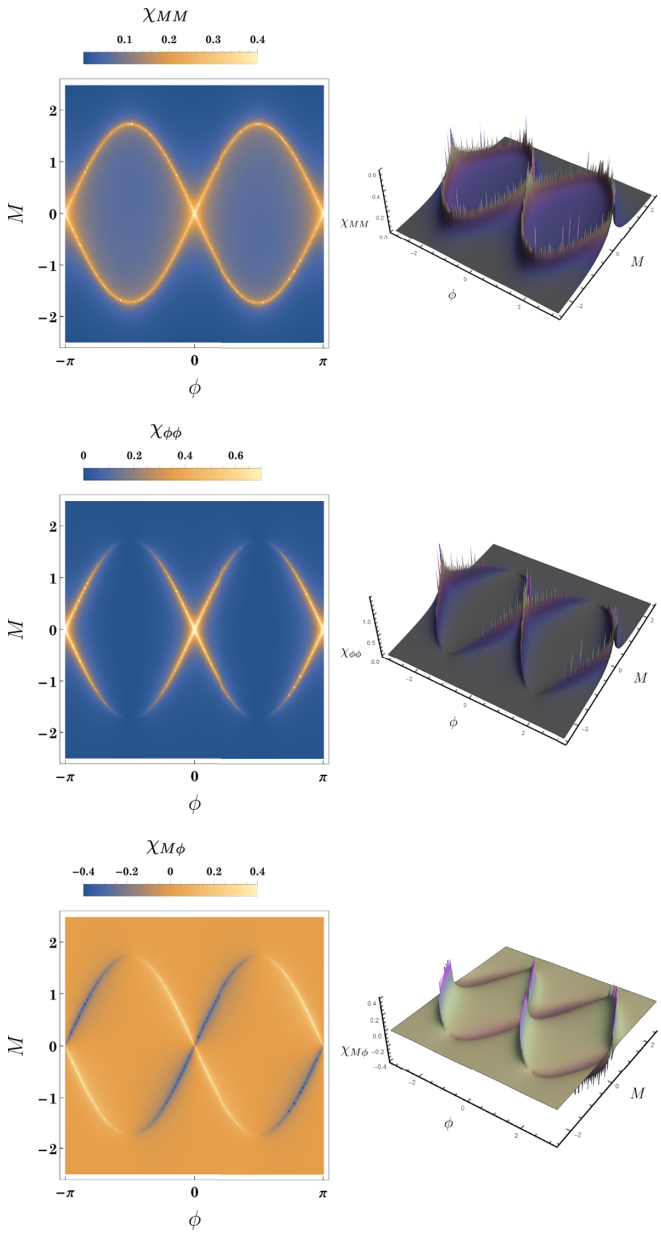


FIG. 2. Fidelity susceptibilities $\chi_{MM}(\phi, M)$ (top), $\chi_{\phi\phi}(\phi, M)$ (middle), and $\chi_{M\phi}(\phi, M)$ (bottom) calculated at temperature $T = 2 \times 10^{-5}$ and $N_x = N_y = 30$ sites.

we did in the Appendix for the (modified) XY model. The analysis, however, is similar but more cumbersome. Since our intent is not to solve particular models, but rather to communicate the idea that tangent directions to critical manifolds are special, we decided to leave these technical details to future work.

VI. CONCLUSIONS

In this paper, we analyzed the thermodynamical limit of the quantum metric along critical submanifolds of theory space. We related its singular behavior to normal directions, which are naturally associated with relevant operators. In the paradigmatic examples of the XY and Haldane models, we have seen (exactly for the first model and numerically for the second) that the normal directions to the critical submanifolds are precisely those where the metric has singular behavior in the thermodynamical limit, while the tangent ones vanish. In both of these models, the critical regions consist of critical submanifolds, which intersect transversally, and the intersection points behave as isolated critical points with enhanced distinguishability. Further analysis of this phenomenon can serve as a future line of research. As we have seen, our theory predicts that there can also be directions associated with relevant operators that lie in the kernel of g_*, a remark which also deserves further investigation. It would be interesting to understand how the above results generalize to the finite-temperature case, where Eq. (18) relating the metric and two-point correlation functions gets modified for the case of the Bures metric. Unlike the Bures metric, for the case of the interferometric metric considered in Ref. [52], the same functional form as (18) is retained—with the expectation value now taken with respect to the appropriate density matrix, and it would be interesting to understand, within the present context, what the physical relevance of considering either metric is. Additionally, the Bogoliubov-Kubo-Mori (BKM) Fisher metric [9,53] seems to be another relevant physical metric in the context of finite-temperature systems, and it would be interesting to understand this case as well.

ACKNOWLEDGMENTS

B.M. and N.P.'s work was partially supported by SQIG – Security and Quantum Information Group of Instituto de Telecomunicações, by Programme (COMPETE 2020) of the Portugal 2020 framework [Project Q.DOT No. 039728 (POCI-01-0247-FEDER-039728)] and the Fundação para a Ciência e a Tecnologia (FCT) through national funds, by FEDER, COMPETE 2020, and by Regional Operational Program of Lisbon, under UIDB/50008/2020 (actions QuRUNNER, QUESTS), Project QuantumMining POCI-01-0145-FEDER-031826 and Project PREDICT PTDC/CCI-CIF/29877/2017. N.P. acknowledges FCT Projects CERN/FIS-PAR/0023/2019, QuantumPrime PTDC/EEI-TEL/8017/2020, and FCT Estímulo ao Emprego Científico Grant No. CEECIND/04594/2017/CP1393/CT000. This work was also supported through Grant No. UID/CTM/04540/2019.

APPENDIX: THERMODYNAMIC LIMIT OF THE QUANTUM METRIC

In this Appendix, we analyze the thermodynamical limit of the quantum metric in the case of the anisotropic XY spin chain in an external magnetic field. A straightforward calculation shows that the components of the quantum metric are given by

$$g_{\lambda\lambda} = \frac{1}{4}\sum_k \frac{\gamma^2 \sin^2(k)}{[(\lambda - \cos(k))^2 + \gamma^2 \sin^2 k]^2},$$

$$g_{\gamma\gamma} = \frac{1}{4}\sum_k \frac{(\lambda - \cos(k))^2 \sin^2(k)}{[(\lambda - \cos(k))^2 + \gamma^2 \sin^2 k]^2},$$

$$g_{\lambda\gamma} = \frac{1}{4}\sum_k \frac{\gamma \sin^2(k)(\cos(k)-\lambda)}{[\gamma^2\sin^2(k)+(\lambda-\cos(k))^2]^2}.$$

Below, we analyze the large-N behavior of the above expressions in the critical region.

1. Segment $\gamma = 0$ and $|\lambda| \leqslant 1$

For the critical submanifold defined by $\gamma = 0$ and $|\lambda| \leqslant 1$, with tangent vector $\frac{\partial}{\partial\lambda}$, we immediately see that $g_{\lambda\lambda} = 0$ independently of N. For $g_{\gamma\gamma}$ the metric becomes

$$g_{\gamma\gamma}(\lambda, 0) = \frac{1}{4}\sum_k \frac{\sin^2(k)}{(\lambda-\cos(k))^2}.$$

For $g_{\gamma\gamma}$ the metric is finite for $|\lambda| > 1$, but for $|\lambda| \leqslant 1$ we have to be careful because the equation

$$\cos(k) = \lambda$$

will have exactly one solution in the thermodynamical limit. In that case, let k_* be the solution. If k is an allowed momentum in a small neighborhood of k_*, i.e., $k = k_* + \delta q$, then we may write $\cos(k) \approx \lambda - \sin(k_*)\delta k = \lambda - \sqrt{1-\lambda^2}\delta k$ and $\sin(k) \approx \sin(k_*) + \lambda\delta k = \sqrt{1-\lambda^2} + \lambda\delta k$ (where the sine is positive, since $k \in [0, \pi]$). Then,

$$\frac{\sin^2(k)}{(\lambda-\cos(k))^2} \sim \frac{(1-\lambda^2)}{(1-\lambda^2)\delta k^2} \sim N^2,$$

so we see that

$$g_{\gamma\gamma}(\lambda, 0) \sim N^2, \quad \text{as } N \to \infty, \text{ for } |\lambda| < 1.$$

For $|\lambda| = 1$, then $\cos(k) \approx \pm 1 + O(\delta k^2)$ and $\sin(k) \approx -\delta k$ (with $\delta k < 0$ since the sine is positive). Hence

$$\frac{\sin^2(k)}{(\lambda-\cos(k))^2} \sim \frac{\delta k^2}{\delta k^4} \sim \frac{1}{\delta k^2} \sim N^2,$$

as before.

2. Line $\lambda = 1$

For the critical submanifold defined by $\lambda = 1$, with tangent vector $\frac{\partial}{\partial\gamma}$, we have

$$g_{\lambda\lambda}(1, \gamma) = \frac{1}{4}\sum_k \frac{\gamma^2\sin^2(k)}{[(1-\cos(k))^2+\gamma^2\sin^2 k]^2}.$$

We have to look at momenta in the neighborhood of $k = 0$, i.e., $k = \delta k$ (with $\delta k > 0$), and then

$$\sin(k) \approx \delta k \quad \text{and} \quad \cos(k) \approx 1 - \tfrac{1}{2}\delta k^2,$$

so

$$\frac{\gamma^2\sin^2(k)}{[(1-\cos(k))^2+\gamma^2\sin^2 k]^2} \approx \frac{\gamma^2\delta k^2}{\left[\frac{1}{4}\delta k^4+\gamma^2\delta k^2\right]^2} \approx \frac{1}{\gamma^2\delta k^2} \sim \frac{N^2}{\gamma^2};$$

thus

$$g_{\lambda\lambda}(1, \gamma) \sim \frac{N^2}{\gamma^2}.$$

This means that as $N \to \infty$, it blows up. Now,

$$g_{\gamma\gamma}(1, \gamma) = \frac{1}{4}\sum_k \frac{(1-\cos(k))^2\sin^2(k)}{[(1-\cos(k))^2+\gamma^2\sin^2 k]^2}.$$

Looking at the momenta in the neighborhood of $k = 0$ as before, we see that

$$\frac{(1-\cos(k))^2\sin^2(k)}{[(1-\cos(k))^2+\gamma^2\sin^2 k]^2} \approx \frac{1}{4}\frac{\delta k^6}{\left(\frac{1}{4}\delta k^4+\gamma^2\delta k^2\right)^2} \approx \frac{1}{4\gamma^4}\delta k^2,$$

so $g_{\gamma\gamma}(1, \gamma)$ is finite for $\gamma \neq 0$ and $g_{\gamma\gamma}(1, \gamma) \sim \frac{1}{N^2\gamma^4}$ as $\gamma \to 0$.

3. Line $\lambda = -1$

For the critical submanifold defined by $\lambda = -1$, with tangent vector $\frac{\partial}{\partial\gamma}$, we have

$$g_{\lambda\lambda}(-1, \gamma) = \frac{1}{4}\sum_k \frac{\gamma^2\sin^2(k)}{[(1+\cos(k))^2+\gamma^2\sin^2 k]^2}.$$

We have to look at momenta in the neighborhood of $k = \pi$, i.e., $k = \pi + \delta k$ (with $\delta k < 0$), and then

$$\sin(k) \approx -\delta k \quad \text{and} \quad \cos(k) \approx -1 + \tfrac{1}{2}\delta k^2,$$

so

$$\frac{\gamma^2\sin^2(k)}{[(1+\cos(k))^2+\gamma^2\sin^2 k]^2} \approx \frac{\gamma^2\delta k^2}{\left[\frac{1}{4}\delta k^4+\gamma^2\delta k^2\right]^2} \approx \frac{1}{\gamma^2\delta k^2} \sim \frac{N^2}{\gamma^2};$$

thus

$$g_{\lambda\lambda}(1, \gamma) \sim \frac{N^2}{\gamma^2}.$$

This means that as $N \to \infty$, it blows up. Now,

$$g_{\gamma\gamma}(-1, \gamma) = \frac{1}{4}\sum_k \frac{(1+\cos(k))^2\sin^2(k)}{[(1+\cos(k))^2+\gamma^2\sin^2 k]^2}.$$

Looking at the momenta in the neighborhood of $k = \pi$ as before, we see that

$$\frac{(1+\cos(k))^2\sin^2(k)}{[(1+\cos(k))^2+\gamma^2\sin^2 k]^2} \approx \frac{1}{4}\frac{\delta k^6}{\left(\frac{1}{4}\delta k^4+\gamma^2\delta k^2\right)^2} \approx \frac{1}{4\gamma^4}\delta k^2,$$

so $g_{\gamma\gamma}(-1, \gamma)$ is finite for $\gamma \neq 0$ and $g_{\gamma\gamma}(-1, \gamma) \sim \frac{1}{N^2\gamma^4}$ as $\gamma \to 0$.

4. Summary

(i) For critical line $\gamma = 0$ and tangent vector $\partial/\partial\lambda$,

$$g_{\lambda\lambda}(\lambda, 0) = 0 \quad \text{and} \quad g_{\gamma\gamma}(\lambda, 0) = \begin{cases} \sim N^2 & \text{as } N \to \infty, \text{ for } |\lambda| \leqslant 1 \\ \text{finite} & \text{otherwise.} \end{cases}$$

(ii) For critical line $\lambda = 1$ and tangent vector $\partial/\partial\gamma$,

$$g_{\lambda\lambda}(1, \gamma) \sim \frac{N^2}{\gamma^2} \quad \text{and} \quad g_{\gamma\gamma}(1, \gamma) = \begin{cases} \text{finite} & \text{for } \gamma \neq 0 \\ \sim \frac{1}{N^2\gamma^4} & \text{as } \gamma \to 0. \end{cases}$$

(iii) For critical line $\lambda = -1$ and tangent vector $\partial/\partial\gamma$,

$$g_{\lambda\lambda}(-1, \gamma) \sim \frac{N^2}{\gamma^2} \quad \text{and} \quad g_{\gamma\gamma}(-1, \gamma) = \begin{cases} \text{finite} & \text{for } \gamma \neq 0 \\ \sim \frac{1}{N^2\gamma^4} & \text{as } \gamma \to 0. \end{cases}$$

In this case we see that $g_* = \lim_{N\to\infty} \frac{g}{N}$ restricted to the critical lines vanishes precisely on tangent vectors to the critical lines, except on the intersections $(\lambda, \gamma) = (\pm 1, 0)$, where it blows up in all directions. On the complementary subspaces to the tangent spaces it always blows up. In this case there are no operators which are not sufficiently relevant.

[1] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of Mathematical Monographs Vol. 191 (American Mathematical Society, Providence, RI, 2000).
[2] R. Balian, Y. Alhassid, and H. Reinhardt, Phys. Rep. **131**, 1 (1986).
[3] B. Mera, P. Mateus, and A. M. Carvalho, IEEE Trans. Inf. Theory **68**, 5619 (2022).
[4] P. Zanardi and N. Paunković, Phys. Rev. E **74**, 031123 (2006).
[5] L. Campos Venuti and P. Zanardi, Phys. Rev. Lett. **99**, 095701 (2007).
[6] P. Zanardi, L. Campos Venuti, and P. Giorda, Phys. Rev. A **76**, 062318 (2007).
[7] A. T. Rezakhani, D. F. Abasto, D. A. Lidar, and P. Zanardi, Phys. Rev. A **82**, 012321 (2010).
[8] N. Paunković and V. Rocha Vieira, Phys. Rev. E **77**, 011129 (2008).
[9] D. Petz, *Quantum Information Theory and Quantum Statistics* (Springer, Berlin, 2007).
[10] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2017).
[11] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).
[12] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[13] A. Fujiwara and H. Nagaoka, Phys. Lett. A **201**, 119 (1995).
[14] D. C. Brody and L. P. Hughston, Proc. R. Soc. London, Ser. A **454**, 2445 (1998).
[15] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, J. Phys. A: Math. Theor. **53**, 023001 (2020).
[16] J. S. Sidhu and P. Kok, AVS Quantum Sci. **2**, 014701 (2020).
[17] J. Suzuki, Y. Yang, and M. Hayashi, J. Phys. A: Math. Theor. **53**, 453001 (2020).
[18] J. Provost and G. Vallee, Commun. Math. Phys. **76**, 289 (1980).
[19] B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Phys. Rev. Lett. **119**, 015702 (2017).
[20] B. Mera, C. Vlachou, N. Paunković, V. R. Vieira, and O. Viyuela, Phys. Rev. B **97**, 094110 (2018).
[21] S. T. Amin, B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Phys. Rev. B **98**, 245141 (2018).
[22] S. Peotta and P. Törmä, Nat. Commun. **6**, 8944 (2015).
[23] A. Julku, S. Peotta, T. I. Vanhala, D.-H. Kim, and P. Törmä, Phys. Rev. Lett. **117**, 045303 (2016).
[24] L. Liang, T. I. Vanhala, S. Peotta, T. Siro, A. Harju, and P. Törmä, Phys. Rev. B **95**, 024515 (2017).
[25] M. Iskin, Phys. Rev. A **97**, 033625 (2018).
[26] S. A. Parameswaran, R. Roy, and S. L. Sondhi, C. R. Phys. **14**, 816 (2013).
[27] R. Roy, Phys. Rev. B **90**, 165139 (2014).
[28] M. Claassen, C. H. Lee, R. Thomale, X.-L. Qi, and T. P. Devereaux, Phys. Rev. Lett. **114**, 236802 (2015).
[29] T. S. Jackson, G. Möller, and R. Roy, Nat. Commun. **6**, 8629 (2015).
[30] C. H. Lee, M. Claassen, and R. Thomale, Phys. Rev. B **96**, 165150 (2017).
[31] J. Wang, J. Cano, A. J. Millis, Z. Liu, and B. Yang, Phys. Rev. Lett. **127**, 246403 (2021).
[32] G. E. Topp, C. J. Eckhardt, D. M. Kennes, M. A. Sentef, and P. Törmä, Phys. Rev. B **104**, 064306 (2021).
[33] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).
[34] T. Ozawa and N. Goldman, Phys. Rev. B **97**, 201117(R) (2018).
[35] G. Palumbo and N. Goldman, Phys. Rev. B **99**, 045154 (2019).
[36] G. Salerno, N. Goldman, and G. Palumbo, Phys. Rev. Res. **2**, 013224 (2020).
[37] T. Ozawa and N. Goldman, Phys. Rev. Res. **1**, 032019(R) (2019).
[38] B. Mera, Phys. Rev. B **101**, 115128 (2020).
[39] T. Ozawa and B. Mera, Phys. Rev. B **104**, 045103 (2021).
[40] B. Mera and T. Ozawa, Phys. Rev. B **104**, 045104 (2021).
[41] B. Mera and T. Ozawa, Phys. Rev. B **104**, 115160 (2021).
[42] J. Wang and Z. Liu, Phys. Rev. Lett. **128**, 176403 (2022).
[43] P. J. Ledwith, A. Vishwanath, and E. Khalaf, Phys. Rev. Lett. **128**, 176404 (2022).
[44] P. Törmä, S. Peotta, and B. A. Bernevig, Nat. Rev. Phys. **4**, 528 (2022).

[45] D. Parker, P. Ledwith, E. Khalaf, T. Soejima, J. Hauschild, Y. Xie, A. Pierce, M. P. Zaletel, A. Yacoby, and A. Vishwanath, arXiv:2112.13837.
[46] C. Northe, G. Palumbo, J. Sturm, C. Tutschku, and E. M. Hankiewicz, Phys. Rev. B **105**, 155410 (2022).
[47] B. Mera, A. Zhang, and N. Goldman, SciPost Phys. **12**, 018 (2022).
[48] D. Tong, Lectures on statistical field theory, 2017, https://www.damtp.cam.ac.uk/user/tong/sft.html.
[49] M. Miyaji, T. Numasawa, N. Shiba, T. Takayanagi, and K. Watanabe, Phys. Rev. Lett. **115**, 261602 (2015).
[50] H. Araki and T. Matsui, Commun. Math. Phys. **101**, 213 (1985).
[51] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
[52] H. Silva, B. Mera, and N. Paunković, Phys. Rev. B **103**, 085127 (2021).
[53] T. Shitara and M. Ueda, Phys. Rev. A **94**, 062316 (2016).

Interferometric geometry from symmetry-broken Uhlmann gauge group with applications to topological phase transitions

Hector Silva,[1,*] Bruno Mera,[2,†] and Nikola Paunković[3,‡]

[1]*Departmento de Física, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal*

[2]*Instituto de Telecomunicações, 1049-001 Lisboa, Portugal;*
Departmento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal;
and Departmento de Física, Instituto Superior Tácnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal

[3]*Instituto de Telecomunicações, 1049-001 Lisboa, Portugal*
and Departmento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal

(Received 13 October 2020; accepted 8 February 2021; published 19 February 2021)

We provide a natural generalization of a Riemannian structure, i.e., a metric, recently introduced by Sjöqvist for the space of nondegenerate density matrices, to the degenerate case, i.e., the case in which the eigenspaces have dimension greater than or equal to 1. We present a physical interpretation of the metric in terms of an interferometric measurement. We apply this metric, physically interpreted as an interferometric susceptibility, to the study of topological phase transitions at finite temperatures for band insulators. We compare the behaviors of this susceptibility and the one coming from the well-known Bures metric, showing them to be dramatically different. While both infer zero-temperature phase transitions, only the former predicts finite-temperature phase transitions as well. The difference in behaviors can be traced back to a symmetry-breaking mechanism, akin to Landau-Ginzburg theory, by which the Uhlmann gauge group is broken down to a subgroup determined by the type of the system's density matrix (i.e., the ranks of its spectral projectors).

DOI: 10.1103/PhysRevB.103.085127

I. INTRODUCTION

Recent advances in information geometry have provided new methods for studying quantum matter and describing macroscopic critical phenomena based on quantum effects. Topological phases of matter are described in terms of *global* topological invariants that are robust against continuous perturbations of the system. An example of these invariants is the Thouless–Kohmoto–Nightingale–den Nijs (TKNN) invariant, mathematically a Chern number associated with the vector bundle of occupied Bloch states over the Brillouin zone. This invariant captures topological phases of matter that could not be understood previously, such as the case of the anomalous Hall insulator [1], which falls into the class of Chern insulators. The classification of topological phases of gapped free fermions is encoded in the so-called periodic table of topological insulators and superconductors [2]. However, by now we know that these phases of matter were just the tip of an iceberg (see Refs. [3–6]). The theory underlying topological phases constitutes a change of paradigm with respect to the Landau theory of phase transitions [7]. The latter is described by means of a *local* order parameter, within the framework of the *symmetry-breaking* mechanism.

One can study phases of matter and the associated phase transitions (in particular, topological ones) through a Riemannian metric on the space of quantum states. One such commonly used structure is based on the notion of fidelity, which is an information theoretical quantity that measures the distinguishability between quantum states. It has been widely used in the study of phase transitions [8–18], since its nonanalytic behavior signals phase transitions.

Note that the mentioned topological invariants, being functions of the Hamiltonian only and not the temperature, characterize topological features at zero temperature. Therefore it is crucial to understand the effect of temperature on topological phase transitions, especially with regard to applications to quantum computers, such as those involving Majorana modes in topological superconductors [19]. To approach this problem, the fidelity, the associated Bures metric, and, in addition, the Uhlmann connection, the generalization of the Berry connection to the case of mixed states, have been probed for systems that exhibit zero-temperature symmetry-protected topological phases [20–24].

Within the context of dynamical phase transitions, occurring when one performs a quench on a system, the information geometric methods based on state distinguishability were applied [25–28]. In particular, for finite-temperature studies, besides the standard notion of fidelity-induced Loschmidt echo, a notion of *interferometric Loschmidt echo* based on the interferometric phase introduced by Sjöqvist *et al.* [29], was also considered. With regard to the associated infinitesimal counterparts, i.e., Riemannian metrics, their behavior is significantly different.

For two-band Chern insulators, the fidelity susceptibility, one of the components of the Bures metric, was considered in detail in Ref. [22]. In particular, it was rigorously proven that the thermodynamic and zero-temperature limits do not

*hectorsilva@tecnico.ulisboa.pt
†bruno.mera@tecnico.ulisboa.pt
‡npaunkov@math.tecnico.ulisboa.pt

 ©2021 American Physical Society

commute: The Bures metric is regular in the thermodynamic limit as one approaches the zero-temperature limit.

In this paper, we provide, through what is called the Ehresmann connection, a natural generalization of a Riemannian structure over the space of nondegenerate density matrices, introduced by Sjöqvist [30], to the degenerate case. Our natural construction reveals a symmetry-breaking mechanism by reducing the gauge group of the Uhlmann principal bundle [31] to a smaller subgroup preserving the *type* of the density matrix, i.e., the ranks of its spectral projectors (see Sec. II for details). This symmetry-breaking mechanism explains the natural enhanced distinguishability provided by the interferometric Riemannian metric. Introducing the notion of a generalized purification, we naturally generalize Sjöqvist's result to the case of degenerate density matrices; see Sec. III. In Sec. IV, we discuss an interferometric measurement probing the Riemannian metric derived. In Sec. V, we apply the derived metric to study finite-temperature phase transitions in the context of band insulators. We present results for this metric in the case of the massive Dirac model, a Chern insulator, in two spatial dimensions and compare them with those obtained using the Bures metric. Our analysis of equilibrium phase transitions is shown to be consistent with the previous study of dynamical phase transitions: The interferometric metric is more sensitive to the change in the parameters than the Bures one. Finally, we present conclusions in Sec. VI.

II. THE GEOMETRY OF THE SJÖQVIST METRIC AND NATURAL GENERALIZATIONS TO DEGENERATE CASES

Consider a quantum system with the corresponding n-dimensional Hilbert space $\mathcal{H}$. Its general mixed state (density matrix) ρ can be, using the spectral decomposition, written as

$$\rho = \sum_{i=0}^{k} p_i P_i, \tag{1}$$

where the real eigenvalues satisfy $p_0 = 0$ and $(i \neq j \Rightarrow p_i \neq p_j)$, while the orthogonal projectors satisfy $(i > 0 \Rightarrow \mathrm{Tr}\, P_i \equiv r_i > 0)$, and $\sum_{i=1}^{k} r_i = r$. We call $r \in \{1, \ldots, n\}$ the *rank* of the state. Note that we do not require for the kernel of ρ to be nontrivial (i.e., $r_0 \equiv \mathrm{Tr}\, P_0 \geqslant 0$), while all other eigenspaces, $\mathcal{H}_i$, are at least one dimensional (such that $\mathcal{H} = \oplus_{i=0}^{k} \mathcal{H}_i$). We call the k-tuple $\tau \equiv (r_1, r_2, \ldots, r_k) \in \mathcal{T}$, with $k \in \{1, \ldots, n\}$ and $(1 \leqslant r_1 \leqslant r_2 \leqslant \cdots \leqslant r_k)$, the *type* of the state ρ, where $\mathcal{T}$ is the set of all possible types. Note that as a consequence of the normalization of density matrices we have the additional constraint

$$\sum_{i=1}^{k} r_i p_i = 1. \tag{2}$$

Consider the set of all density operators of type τ, denoted by B_τ. The union, over the types $\tau \in \mathcal{T}$, of all sets B_τ forms the set of all possible states of a given system,

$$\begin{aligned} B &= \bigcup_{\tau \in \mathcal{T}} B_\tau \\ &= \{\rho \in \mathcal{H} \otimes \mathcal{H}^* : \rho^\dagger = \rho \text{ and } \rho \geqslant 0 \text{ and } \mathrm{Tr}\, \rho = 1\}. \end{aligned} \tag{3}$$

We would like to analyze the geometry of the B_τ's and see whether it is possible to induce a Riemannian metric on them along the lines of the metric introduced by Sjöqvist [30], for the case of type $\tau = (1, 1, \ldots, 1)$, for some $r = k$. We will do so by introducing gauge-invariant Riemannian metrics and associated Ehresmann connections in suitably chosen principal bundles P_τ with corresponding base spaces B_τ. Observe that every state ρ is completely specified in terms of its "*classical part*," the vector of probabilities $\sqrt{\mathbf{p}} = (\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_k})$ satisfying the normalization constraint (2), and its "*quantum part*," the mutually orthogonal projectors $P_1, P_2, \ldots, P_k$ (note that P_0 is then determined unambiguously, $P_0 = I - \sum_{i=1}^{k} P_i$), which we compactly denote by $\mathbf{P} = (P_1, P_2, \ldots, P_k)$. We will explore a particular gauge degree of freedom in describing the quantum part in our construction. Namely, each eigenspace projector P_i is uniquely specified by an orthonormal basis $\beta_i = \{|e_{i,j}\rangle : j = 1, \ldots, r_i\}$. However, the basis β_i itself is not uniquely determined by P_i. Indeed, every basis $U\beta_i = \{U|e_{i,j}\rangle : j = 1, \ldots, r_i\}$ with U being a unitary that acts nontrivially only on the image of P_i, the subspace $\mathcal{H}_i$, defines the same projector P_i.

We then define (the total space of) a principal bundle P_τ as the set of all k-tuples of pairs $p_\tau = ((p_i, \beta_i))_{i=1}^{k}$, such that $(\sqrt{\mathbf{p}}, \mathbf{P})$ give rise to well-defined type τ density operators (observe that $p_i \neq p_j$ for all $i \neq j$). This space comes equipped with an obvious projection to the base space B_τ and is given by

$$\pi_\tau(p_\tau) \equiv \sum_{i=1}^{k} p_i P_i = \rho, \tag{4}$$

with the fibers being isomorphic to the product of the corresponding unitary groups in the type τ,

$$G_\tau \equiv \prod_{i=1}^{k} \mathrm{U}(r_i). \tag{5}$$

The group G_τ acts on the right in the obvious way; for $U_i \in \mathrm{U}(r_i)$, we write $U_i = [(U_i)_j^{j'}]_{1 \leqslant j, j' \leqslant r_i} \in \mathrm{U}(r_i)$, and then $\beta_i \cdot U_i$ is given by

$$|e_{i,j}\rangle \mapsto \sum_{j'=1}^{r_i} |e_{i,j'}\rangle (U_i)_j^{j'}, \quad j = 1, \ldots, r_i. \tag{6}$$

By introducing *generalized amplitudes* $w_i \in \mathbb{C}^{n \times r_i}$ as matrices whose columns are vectors $|e_{i,j}\rangle \in \mathbb{C}^n$, $j = 1, \ldots, r_i$, i.e., $w_i \equiv (|e_{i,1}\rangle\ |e_{i,2}\rangle \cdots |e_{i,r_i}\rangle)$, $i = 1, \ldots, k$, we can see P_τ as

$$\begin{aligned} P_\tau = \Bigg\{ &((p_i, w_i))_{i=1}^{k} : \sum_{i=1}^{k} p_i w_i w_i^\dagger \in B_\tau \\ &\text{and } w_i^\dagger w_i = I_{r_i}, \text{ for all } i = 1, \ldots, k, \\ &\text{and } p_i \neq p_j, \text{ for all } i \neq j \Bigg\}, \end{aligned} \tag{7}$$

and the right action of the gauge group is given by $w_i \mapsto w_i \cdot U_i$, with $U_i \in \mathrm{U}(r_i)$. With this notation, we finally introduce a suitable "Hermitian form" (note that it is not a scalar product, as P_τ is not a linear space), which will define horizontal

subspaces, by the formula

$$\langle p_\tau, p'_\tau\rangle_\tau \equiv \sum_{i=1}^{k}\sqrt{p_i p'_i}\,\mathrm{Tr}(w_i^\dagger w'_i)$$
$$= \sum_{i=1}^{k}\mathrm{Tr}[(\sqrt{p_i}w_i^\dagger)(\sqrt{p'_i}w'_i)]. \quad (8)$$

Observe that it is clear that this pairing arises from the restriction of the usual Hermitian inner product in $\bigoplus_{i=1}^{k}\mathbb{C}^{n\times r_i} \cong \mathbb{C}^{n\times r}$.

Additionally, this allows for a convenient comparison with the Uhlmann principal bundle

$$P_r^{\mathrm{Uh}} = \{w \in \mathbb{C}^{n\times r} : \pi(w) \equiv ww^\dagger = \rho \in B, \text{ with } \mathrm{rank}(\rho) = r\}, \quad (9)$$

where the typical fiber is $\mathrm{U}(r) \subset \mathbb{C}^{r\times r}$, whose elements act from the right ($w \mapsto w\cdot U$), and the Hermitian form, induced by the Hilbert-Schmidt scalar product on the space of linear operators from $\mathbb{C}^{r\times r}$, is

$$\langle w, w'\rangle = \mathrm{Tr}(w^\dagger w'). \quad (10)$$

Note that the base space for the Uhlmann bundle is the set of density matrices with rank r, which is the union of all B_τ sharing the same rank. Observe that for one such τ, P_τ can be identified as a subset of P_r^{Uh}. This follows from the map

$$P_\tau \ni ((p_i, w_i))_{i=1}^{k} \mapsto (\sqrt{p_1}w_1, \ldots, \sqrt{p_k}w_k) \in \bigoplus_{i=1}^{k}\mathbb{C}^{n\times r_i}, \quad (11)$$

being an embedding of P_τ. Moreover, once we identify $\bigoplus_{i=1}^{k}\mathbb{C}^{n\times r_i} \cong \mathbb{C}^{n\times r}$, the image sits precisely in P_r^{Uh}. In other words $P_\tau \subset P_r^{\mathrm{Uh}}$, and also π_τ equals the restriction of the projection of the Uhlmann bundle to P_τ ($p_i \neq p_j$, for all $i \neq j$, guarantees this), the image being precisely B_τ. We remark that the gauge group of the Uhlmann bundle is far larger than the one for the principal bundle $P_\tau \to B_\tau$. By passing to a preferred type, we performed a symmetry-breaking operation from $\mathrm{U}(r)$ to $G_\tau = \prod_{i=1}^{k}\mathrm{U}(r_i) \subset \mathrm{U}(r)$. This is another way to see why interferometriclike quantities, such as the interferometric Loschmidt echo, in certain applications develop nonanalyticities, while the ones based on the fidelity do not (see, for example, Ref. and the references therein): The former have smaller space to "go through," while the latter can, following the "broader" Uhlmann connection, instead of the interferometric ones, avoid possible sources of nonanalyticities.

III. DISTANCE MEASURES AND RIEMANNIAN METRICS

Consider now two points, $p_\tau = ((p_i, w_i))_{i=1}^{k}$ and $q_\tau = ((q_i, v_i))_{i=1}^{k} \in P_\tau$. By making use of Eq. (8) one can define a distance between elements p_τ and q_τ in the total space of the principal bundle given by

$$d_\tau^2(p_\tau, q_\tau) = 2(1 - \mathrm{Re}(\langle p_\tau, q_\tau\rangle_\tau))$$
$$= 2\left(1 - \sum_{i=1}^{k}\sqrt{p_i q_i}\ \mathrm{Re}\left(\mathrm{Tr}(w_i^\dagger v_i)\right)\right). \quad (12)$$

The fact that d_τ is a distance follows from the fact that it is the restriction of the usual distance in $\oplus_{i=1}^{k}\mathbb{C}^{n\times r_i}$, where we see P_τ as a subset of this space through the map of Eq. (11). One can use this distance to define a distance on B_τ, through the formula

$$d_I^2(\rho,\sigma) = \inf\left\{d_\tau^2(p_\tau, q_\tau) : \pi(p_\tau) = \rho \text{ and } \pi(q_\tau) = \sigma, \text{ for } p_\tau, q_\tau \in P_\tau\right\}. \quad (13)$$

The associated infinitesimal counterparts of the distances defined above are Riemannian metrics on P_τ and B_τ, respectively. The Riemannian metric on P_τ, which is gauge invariant, allows for the definition of what is called an Ehresmann connection over P_τ, and this, in turn, defines a metric downstairs over the base space B_τ.

Another way to see that $d_\tau^2(p_\tau, q_\tau)$ is indeed a metric is through what we call "*generalized purifications*." Let us introduce "*ancilla*" amplitudes $\mathrm{w}_i \in \mathbb{C}^{k\times 1}$, with $i = 1, 2, \ldots, k$, such that $\mathrm{w}_i\mathrm{w}_i^\dagger = \mathrm{P}_i \in \mathbb{C}^{n\times n}$ are *fixed* orthogonal projectors of rank 1 (i.e., P_i do not depend on the choice of the state), satisfying $\mathrm{P}_i\mathrm{P}_j = \delta_{ij}I_k$ and $\sum_{i=1}^{k}\mathrm{P}_i = I_k$. Define a generalized purification of state ρ, associated with the corresponding p_τ, as

$$|p_\tau\rangle = \sum_{i=1}^{k}\sqrt{p_i}w_i \otimes \mathrm{w}_i. \quad (14)$$

Then, we have that the scalar product between $|p_\tau\rangle$ and $|q_\tau\rangle$, induced by the Hilbert-Schmidt scalar product in the corresponding factor spaces, is

$$\langle p_\tau, q_\tau\rangle = \sum_{i,j=1}^{k}\sqrt{p_i q_j}\langle w_i, v_j\rangle\langle \mathrm{w}_i, \mathrm{w}_j\rangle$$
$$= \sum_{i=1}^{k}\sqrt{p_i q_i}\langle w_i, v_i\rangle$$
$$= \sum_{i=1}^{k}\sqrt{p_i q_i}\,\mathrm{Tr}(w_i^\dagger v_i)$$
$$= \langle p_\tau, q_\tau\rangle_\tau, \quad (15)$$

where the second equality is because w_i and w_j are orthogonal for $i \neq j$. Thus the distance $d_\tau(p_\tau, q_\tau)$ is nothing but the standard Hilbert-Schmidt distance between the generalized purifications $|p_\tau\rangle$ and $|q_\tau\rangle$.

As in Eq. (7), if we take the w_i's as (row) vectors $|w_i\rangle = [|e_{i,1}\rangle\ |e_{i,2}\rangle \cdots |e_{i,r_i}\rangle]$ whose entries are (column) vectors $|e_{i,j}\rangle$, one can by analogy generalize the quantum part of the metric for the nondegenerate case, the so-called "*interferometric metric*," which has $r_i = 1$, $i = 1, \ldots, k$,

$$g_I^{\mathrm{Q}} = \sum_{i=1}^{k} p_i\langle dw_i|(I_n - w_i w_i^\dagger)|dw_i\rangle$$
$$= \sum_{i=1}^{k} p_i\langle de_{i,1}|(I_n - |e_{i,1}\rangle\langle e_{i,1}|)|de_{i,1}\rangle, \quad (16)$$

to the degenerate case, in which U(1) degree of freedom of each $w_i = |e_i\rangle$ is replaced by the $\mathrm{U}(r_i)$ degree of freedom of

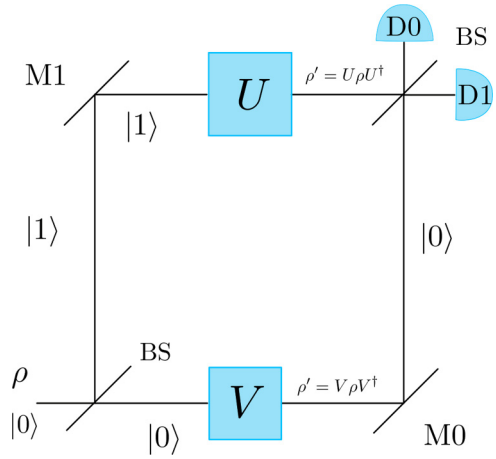


FIG. 1. Interferometric measurement to probe the generalized metric g_I. BS represents beam-splitter, M_0 and M_1 mirrors in arms 0 and 1, respectively, and analogously for detectors D_0 and D_1.

each $w_i = [|e_{i,1}\rangle\ |e_{i,2}\rangle \cdots |e_{i,r_i}\rangle]$,

$$\begin{aligned} g_I^{\mathrm{Q}} &= \sum_{i=1}^{k} p_i \langle dw_i | \big(I_n - w_i w_i^\dagger\big) | dw_i \rangle \\ &= \sum_{i=1}^{k} p_i \langle dw_i | \left[I_n - \left(\sum_{j=1}^{r_i} |e_{i,j}\rangle\langle e_{i,j}| \right) \right] | dw_i \rangle \\ &= \sum_{i=1}^{k} p_i \langle dw_i | \big(I_n - P_i\big) | dw_i \rangle, \end{aligned} \tag{17}$$

with $|dw_i\rangle = [|de_{i,1}\rangle\ |de_{i,2}\rangle \cdots |de_{i,r_i}\rangle]$, $i = 1, \ldots, k$. Indeed, in Appendix A we prove that this intuitive generalization is the correct result describing the infinitesimal counterpart of the distance in Eq. (13).

IV. INTERFEROMETRIC MEASUREMENT INTERPRETATION

Consider the following experiment depicted in Fig. 1. A particle is entering the Mach-Zehnder interferometer from the input arm 0, given by the state $|0\rangle$, with its internal degree of freedom in a mixed state ρ. Both the input and the output beam splitters are balanced, described by the same unitary matrix, say, the one given by $|0\rangle \to (|0\rangle + i|1\rangle)/\sqrt{2}$. In arm 0 a unitary $V = \sum_{i=0}^{k} P_i V P_i$ is applied to the internal degree of freedom; that is, V is the most general unitary that commutes with ρ. In arm 1 a unitary $U = U(\delta t) \in \mathrm{U}(n)$ is applied for a time period δt, changing the state of the internal degree of freedom to $\rho' = U\rho U^\dagger$. The particle is detected at detectors D0 and D1, with the corresponding probabilities pr_0 and pr_1. In our case, we have that $\mathrm{pr}_1 \leqslant \mathrm{pr}_0$, and for $U = V$ we have full constructive interference at the output arm 0, giving $\mathrm{pr}_0 = 1$. In general, we have that

$$\mathrm{pr}_1^{\max} = \max_{\{V_i\}} (\mathrm{pr}_1) = 1 - \tfrac{1}{4} d_I^2(\rho, \rho + \delta\rho), \tag{18}$$

where $d_I^2(\rho, \rho + \delta\rho) \approx g_I(\dot\rho, \dot\rho)\delta t^2$ is the "infinitesimal" distance between ρ and $\rho' = \rho + \delta\rho$, where $\delta\rho = \dot\rho \delta t$ (see Appendix B for a detailed proof). Note that in the case of the Hadamard matrix, given by $|\ell\rangle \to [|0\rangle + (-1)^\ell |1\rangle]/\sqrt{2}$, with $\ell \in \{0, 1\}$, the roles of arms 0 and 1 are exchanged.

V. INTERFEROMETRIC METRIC IN THE CONTEXT OF BAND INSULATORS

Suppose we have a family of band insulators with two bands described by the Hamiltonian

$$\mathcal{H}(M) = \int_{\mathrm{BZ}^d} \frac{d^d k}{(2\pi)^d} \psi_{\mathbf{k}}^\dagger d^\mu(\mathbf{k}; M) \sigma_\mu \psi_{\mathbf{k}}, \tag{19}$$

parametrized by M (M can be some intrinsic parameter, such as the hopping), where σ_μ, $\mu = 1, 2, 3$, are the Pauli matrices, $\mathbf{k}$ is the crystalline momentum in a d-dimensional Brillouin zone BZ^d, with $d = 1, 2, 3$, and $\Psi_{\mathbf{k}}^\dagger$ is an array of two creation operators for fermions at momentum $\mathbf{k}$. We assume that the system is gapped for generic values of M, meaning that the vector $d = (d^1, d^2, d^3)$ is nonvanishing as a function of $\mathbf{k}$. For a certain value of M_c, we assume that the vector has isolated zeros. This assumption is generically correct for the $d = 1, 2$ momenta coordinates plus the mass M, as one needs to tune three parameters for a Hermitian matrix to have two eigenvalues cross.

The pullback of the interferometric metric that we have described in Sec. III,

$$g = \frac{1}{4} \sum_i r_i \frac{dp_i^2}{p_i} + \sum_i p_i \,\mathrm{Tr}\,(P_i dP_i dP_i), \tag{20}$$

with $\rho = \sum_i p_i P_i$ and $\mathrm{Tr}\, P_i = r_i$, by the map induced by the Gibbs state

$$M \mapsto \rho(M) = Z^{-1} \exp[-\beta \mathcal{H}(M)], \tag{21}$$

where Z is the partition function, is given by

$$\begin{aligned} ds^2 = \frac{1}{4} \int_{\mathrm{BZ}^d} \frac{d^d k}{(2\pi)^d} \Bigg\{ & \frac{1}{\cosh(\beta E) + 1} \\ & \times \left[\beta^2 \left(\frac{\partial E}{\partial M} \right)^2 + \cosh(\beta E) \delta_{\mu\nu} \frac{\partial n^\mu}{\partial M} \frac{\partial n^\nu}{\partial M} \right] \Bigg\} dM^2, \end{aligned} \tag{22}$$

where we omitted the obvious dependence on $\mathbf{k}$ and M of the quantities E and n^μ. We provide a technical derivation of this result in Appendix C. This result should be compared with the pullback of the Bures metric for $d = 2$, which yields (see Ref. [22])

$$\begin{aligned} g_{\mathrm{Bures}} = \frac{1}{4} \int_{\mathrm{BZ}^d} \frac{d^d k}{(2\pi)^d} \Bigg[& \frac{1}{\cosh(\beta E) + 1} \beta^2 \left(\frac{\partial E}{\partial M} \right)^2 \\ & + \frac{\cosh(\beta E) - 1}{\cosh(\beta E)} \delta_{\mu\nu} \frac{\partial n^\mu}{\partial M} \frac{\partial n^\nu}{\partial M} \Bigg] dM^2. \end{aligned} \tag{23}$$

The two expressions have dramatically different behaviors, when it comes to taking the zero-temperature limit.

Naively, one would say that both yield the pullback of the Fubini-Study metric, which is the pure-state metric,

$$g_0 = \frac{1}{4} \int_{\mathrm{BZ}^d} \frac{d^d k}{(2\pi)^d} \delta_{\mu\nu} \frac{\partial n^\mu}{\partial M} \frac{\partial n^\nu}{\partial M} dM^2. \tag{24}$$

Note that for gapless points the vector n is not defined and the expression for g_0 becomes (potentially) singular. However, due to the gapless points, the integrands must be carefully analyzed in the neighborhoods of these points, as the singularities can be avoided in some cases. In fact, it was shown that if the gapless points are isolated in momentum space, then an expansion near these points of the integrand function yields a regular result [22]. Namely, because of the inequality

$$\frac{1}{2}\frac{1}{\cosh(x)} < \frac{1}{\cosh(x)+1} < \frac{1}{\cosh(x)} \quad \text{for all } x \in \mathbb{R}, \tag{25}$$

we can write

$$\frac{1}{\cosh(\beta E)+1}\beta^2\left(\frac{\partial E}{\partial M}\right)^2 + \frac{\cosh(\beta E)-1}{\cosh(\beta E)}\delta_{\mu\nu}\frac{\partial n^\mu}{\partial M}\frac{\partial n^\nu}{\partial M} < \frac{1}{\cosh(\beta E)}\left[\beta^2\left(\frac{\partial E}{\partial M}\right)^2 + [\cosh(\beta E)-1]\delta_{\mu\nu}\frac{\partial n^\mu}{\partial M}\frac{\partial n^\nu}{\partial M}\right]. \tag{26}$$

Expansion for small βE yields that up to $\mathrm{O}((\beta E)^4)$ the integrand is upper bounded by

$$\frac{\beta^2}{\cosh(\beta E)}\delta_{\mu\nu}\frac{\partial d^\mu}{\partial M}\frac{\partial d^\nu}{\partial M}, \tag{27}$$

which is regular in the limit $\beta \to \infty$. Hence the potential singularities arising from the gapless region are regularized by the Bures prescription. However, in the case of the interferometric metric, considering the integrand

$$\frac{1}{\cosh(\beta E)+1}\left[\beta^2\left(\frac{\partial E}{\partial M}\right)^2 + \cosh(\beta E)\delta_{\mu\nu}\frac{\partial n^\mu}{\partial M}\frac{\partial n^\nu}{\partial M}\right] \tag{28}$$

near $E = 0$ gives us

$$\frac{1}{\cosh(\beta E)+1}\left[\beta^2\left(\frac{\partial E}{\partial M}\right)^2 + \left(1+\frac{1}{2}\beta^2E^2\right)\delta_{\mu\nu} \times \frac{\partial n^\mu}{\partial M}\frac{\partial n^\nu}{\partial M} + \mathrm{O}((\beta E)^4)\right]. \tag{29}$$

In this case, we cannot get rid of the singular factor

$$\delta_{\mu\nu}\frac{\partial n^\mu}{\partial M}\frac{\partial n^\nu}{\partial M}, \tag{30}$$

which appears once in the second term without the regularizing coefficient β^2E^2 which above allowed for the identification of the regular quantity

$$\beta^2\left(\frac{\partial E}{\partial M}\right)^2 + \beta^2E^2\delta_{\mu\nu}\frac{\partial n^\mu}{\partial M}\frac{\partial n^\nu}{\partial M} = \beta^2\delta_{\mu\nu}\frac{\partial d^\mu}{\partial M}\frac{\partial d^\nu}{\partial M}. \tag{31}$$

This implies that the limit $\beta \to \infty$ yields singular behavior for g, provided the same happens with g_0, but not the other way around; that is, singular behavior on the finite-temperature metric does not imply zero-temperature singular behavior. In other words, while in the case of the Bures metric the thermodynamic and the zero-temperature limits did not commute, in the interferometric case they do, because the singular behavior of the gapless points is recovered, as one considers a small neighborhood of these points and takes the zero-temperature limit. In the following, we will consider the massive Dirac model to illustrate the different behaviors of the two metrics.

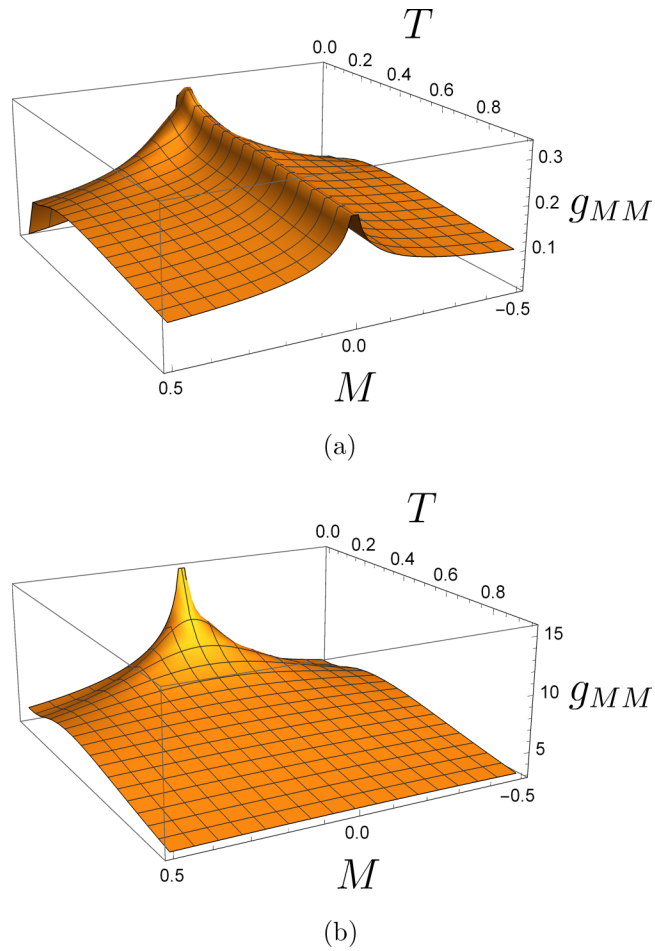


FIG. 2. The different behavior of the metrics with temperature T and the parameter M driving the topological phase transition. (a) Interferometric metric for the massive Dirac model: The topological phase transition is captured for all temperatures. (b) Bures metric for the massive Dirac model: The topological phase transition is captured only at zero temperature.

A. Massive Dirac model

We consider the massive Dirac model, a band insulator in two spatial dimensions, described by Eq. (19), with

$$d(\mathbf{k};M) = (\sin(k_x), \sin(k_y), M - \cos(k_x) - \cos(k_y)), \tag{32}$$

where $\mathbf{k} = (k_x, k_y)$ is the quasimomentum in the two-dimensional Brillouin zone BZ^2 and M is a real parameter. The model exhibits topological phase transitions [32]. We will focus on the one occurring at $M = 0$, where the Chern number goes from $+1$, for $M \to 0^-$, to -1, for $M \to 0^+$. Figure 2 describes the interferometric metric [Fig. 2(a)] and the Bures metric [Fig. 2(b)] in the thermodynamic limit.

As argued above, the Bures metric is regular if one considers the thermodynamic limit and then the zero-temperature limit. The same does not hold for the interferometric metric. In fact, we can see that the interferometric metric knows about the quantum phase transition taking place at $T = 0$ even at finite temperatures. The reason is that in passing from one metric to the other the symmetry was broken, namely, $\mathrm{U}(r) \to \prod_{i=1}^{k}\mathrm{U}(r_i)$, and therefore there is enhanced distin-

guishability. Indeed, in the interferometric case, whenever the gap closes, we expect a phase transition, even at finite temperatures, because then there are states which according to a Boltzmann-Gibbs distribution become degenerate in probability; hence the gap closing changes the type of the density matrix involved. Whether such singular behavior of the interferometric metric is indeed observable for macroscopic many-body systems is an open question. While the straightforward implementation of the interferometric experiment described in Sec. IV seems to be, at least technologically, infeasible, as it would require maintaining Schrödinger-cat-like macroscopic states, possible variations are argued to be able to reveal the singular behavior of the interferometric metric at finite temperatures (see Sec. V of Ref. [33]).

VI. CONCLUSIONS

In this paper, we have generalized Sjöqvist's interferometric metric introduced in Ref. [30] to the degenerate case. For this purpose, we have introduced generalized amplitudes and purifications. We have analyzed an interpretation of the metric in terms of a suitably generalized interferometric measurement, accommodating for the non-Abelian character of our gauge group, as opposed to the Abelian gauge group used in the nondegenerate case. We have applied the induced Riemannian structure, physically interpreted as a susceptibility, to the study of topological phase transitions at finite temperatures for band insulators. To the best of our knowledge, this is the first study of finite-temperature equilibrium phase transitions using interferometric geometry. The inferred critical behavior is very different from that of the Bures metric. The interferometric metric is more sensitive to the change in parameters than the Bures one, and unlike the latter, in addition to zero-temperature phase transitions, infers finite-temperature phase transitions as well. This sensitivity can be traced back to a symmetry-breaking mechanism, much in the same spirit of the Landau-Ginzburg theory. In our case, by fixing the type of the density matrix considered, a gauge group is broken down to a subgroup.

It would be very interesting to analyze the interferometric curvature, an analog of the usual Berry curvature, generalized to this mixed setting, associated with the Ehresmann connection presented in this paper. Since the curvature is intrinsically related to topological phenomena, this analysis might very well unravel new symmetry-protected topological phases in the mixed-state case and potentially help refine the classification of topological matter. It would be also interesting to compare the critical behavior of different many-body systems in terms of interferometric metrics corresponding to different types of density matrices. Recent study of the fidelity susceptibility indicated that its singular behavior around regions of criticality has preferred directions on the parameter space [34]. Performing a similar analysis for the interferometric critical geometry is another possible line of future research. Finally, probing experimentally the introduced interferometric metrics is a relevant topic of future investigation.

ACKNOWLEDGMENTS

B.M. and N.P. are thankful for the support from Security and Quantum Information Group (SQIG), the Instituto de Telecomunicações (IT) Research Unit, Reference No. UIDB/50008/2020, funded by Fundação para a Ciência e a Tecnologia (FCT), European funds, namely, Horizon 2020 project SPARTA, as well as projects QuantMining (No. POCI-01-0145-FEDER-031826) and PREDICT (No. PTDC/CCI-CIF/29877/2017). N.P. acknowledges FCT Estímulo ao Emprego Científico Grant No. CEECIND/04594/2017/CP1393/CT0006.

APPENDIX A: INDUCED RIEMANNIAN METRICS

Let us look again at the principal bundle P_τ, for a fixed type $\tau = (r_1, \ldots, r_k)$. In this case, a point in P_τ is given by $p_\tau = ((p_i, w_i))_{i=1}^k$ and can be equivalently represented as $p_\tau = ((p_i)_{i=1}^k, (w_i)_{i=1}^k))$. With this identification, we can separate p_τ into its "*classical*" and "*quantum*" parts:

(i) A classical probability amplitude vector $\sqrt{\mathbf{p}} = (\sqrt{p_1}, \ldots, \sqrt{p_k})$, with $\sum_{i=1}^k p_i = 1$ and, for each $i \in \{1, \ldots, k\}$, $p_i > 0$. Note that the set of all classical probability amplitudes is in fact contained in the $(k-1)$-dimensional sphere and the associated classical Fisher metric is, up to a factor of $1/4$, the usual round metric in the sphere S^{k-1}.

(ii) A quantum part which is a k-tuple, i.e., a sequence of matrices $(w_1, \ldots, w_k)$, each of them identifying a r_i-unitary frame in $\mathbb{C}^n$, i.e., $w_i \in \mathrm{V}_{r_i}(\mathbb{C}^n)$, where

$$V_{r_i}(\mathbb{C}^n) = \{w_i \in \mathbb{C}^{n\times r_i} : w_i^\dagger w_i = I_k\} \subset \mathbb{C}^{n\times r_i}, \quad i = 1, \ldots, k, \tag{A1}$$

commonly known as the *Stiefel* manifold of r_i-unitary frames in $\mathbb{C}^n$.

Our aim is to compute the Riemannian metric in the base space B_τ for a given type $\tau = (r_1, \ldots, r_k)$. For this purpose, we will first look at the tangent space at a point p_τ, which is isomorphic to the direct sum

$$T_{p_\tau}P_\tau \cong T_{\sqrt{\mathbf{p}}}S^{k-1} \oplus \left(\bigoplus_{i=1}^k T_{w_i}V_{r_i}(\mathbb{C}^n)\right). \tag{A2}$$

This isomorphism follows from the factorization into classical and quantum parts: For every curve in the total space P_τ, there will be a tangent vector for each of the curves induced by projection in the different factors of P_τ.

The classical components have no gauge ambiguity. The quantum components, however, have a $\mathrm{U}(r_i)$ gauge degree of freedom for each matrix w_i, $i = 1, \ldots, k$. This gauge ambiguity corresponds to variations along the fibers, as we will mention later on. From a physical standpoint, the exact point in the fiber has no significance, since the matrices w_i will be projected onto the base space, where the projectors P_i are gauge invariant; namely, w_i and $w_i \cdot U$, for $U \in \mathrm{U}(r_i)$, give rise to the same projector $P_i = w_i w_i^\dagger = w_i U U^\dagger w_i^\dagger$, for all $i = 1, \ldots, k$. Hence we need to define the horizontal subspaces of the tangent spaces to P_τ, in order to uniquely represent the tangent spaces to the base space upstairs, i.e., in the tangent spaces to P_τ. Mathematically, this notion is referred to as an *Ehresmann connection*; see, for example, Sec. 6.3 of Ref. [35].

Before we proceed, let us focus on one of the Stiefel manifolds, say, for a fixed $i \in \{1, \ldots, k\}$, $V_{r_i}(\mathbb{C}^n)$. For convenience, we define the projection onto the space of projectors of rank r_i, identified with the Grassmannian of r_i planes in $\mathbb{C}^n$, i.e., the manifold of linear subspaces of dimension r_i in $\mathbb{C}^n$,

$$\begin{aligned} \pi_i : V_{r_i}(\mathbb{C}^n) &\to \mathrm{Gr}_{r_i}(\mathbb{C}^n) \\ w_i &\mapsto P_i = w_i w_i^\dagger. \end{aligned} \tag{A3}$$

Consider a curve in the Stiefel manifold

$$\gamma_{w_i} : [0, 1] \ni t \mapsto \gamma_{w_i}(t) \in V_{r_i}(\mathbb{C}^n) \tag{A4}$$

subject to the initial conditions $\gamma_{w_i}(0) = w_i$ and $\frac{d\gamma_{w_i}}{dt}|_{t=0} = \dot{w}_i \equiv \widetilde{v}$. From the definition of $V_{r_i}(\mathbb{C}^n)$, the tangent spaces are

$$T_{w_i} V_{r_i}(\mathbb{C}^n) = \{\dot{w}_i \in \mathbb{C}^{n\times r_i} : \dot{w}_i^\dagger w_i + w_i^\dagger \dot{w}_i = 0\}. \tag{A5}$$

The vertical space at $w_i \in V_{r_i}(\mathbb{C}^n)$ is the set of tangent vectors in $T_{w_i} V_{r_i}(\mathbb{C}^n)$, such that its infinitesimal projection onto the base space is zero; that is,

$$\begin{aligned} &\frac{d}{dt}\big(\pi_i(\gamma_{w_i}(t))\big)\big|_{t=0} = 0, \\ \Leftrightarrow\ &\frac{d}{dt}\big(\gamma_{w_i}(t)\gamma_{w_i}^\dagger(t)\big)\big|_{t=0} = \dot{w}_i w_i^\dagger + w_i \dot{w}_i^\dagger = 0. \end{aligned} \tag{A6}$$

The vertical space is then given by

$$V_{w_i} = \big\{\dot{w}_i \in T_{w_i} V_{r_i}(\mathbb{C}^n) : \dot{w}_i w_i^\dagger + w_i \dot{w}_i^\dagger = 0\big\}. \tag{A7}$$

The projection π_i has the derivative $d\pi_i = w_i dw_i^\dagger + dw_i w_i^\dagger$, and the vertical tangent vectors are in the kernel of this linear map. Given a fiber of π_i and a choice of a w_i in this fiber, we can diffeomorphically identify the fiber with $\mathrm{U}(r_i)$ by right multiplication. Suppose we take $X \in \mathfrak{u}(r_i)$, identified as an anti-Hermitian matrix in the usual way, and choose a curve $t \mapsto w_i(t) = w_i \cdot e^{tX}$. Clearly, the projection onto the base is invariant under this transformation:

$$\begin{aligned} w_i(t) w_i^\dagger(t) &= w_i e^{tX} (w_i e^{tX})^\dagger \\ &= w_i e^{tX} e^{-tX} w_i^\dagger = w_i w_i^\dagger. \end{aligned} \tag{A8}$$

The tangent vector to the fiber can now be written as $\frac{dw_i}{dt}|_{t=0} = \dot{w}_i = w_i \cdot X$, which satisfies the condition for vertical matrices

$$\begin{aligned} \dot{w}_i w_i^\dagger + w_i \dot{w}_i^\dagger &= w_i X w_i^\dagger + w_i X^\dagger w_i \\ &= w_i X w_i^\dagger - w_i X w_i^\dagger \\ &= 0. \end{aligned} \tag{A9}$$

Hence, by dimensionality, our vertical space can also be seen as

$$V_{w_i} = \big\{\dot{w}_i \in T_{w_i} V_{r_i}(\mathbb{C}^n) : \dot{w}_i = w_i \cdot X,\ \text{where } X^\dagger = -X\big\}. \tag{A10}$$

We are now in a condition to define the horizontal subspaces, which will simply be the collection of tangent vectors $\dot{w}_i$ that are orthogonal to V_{w_i}

$$\begin{aligned} H_{w_i} &= \big(V_{w_i}\big)^\perp \\ &= \big\{\dot{w} \in T_{w_i} V_{r_i}(\mathbb{C}^n) : \langle \dot{w}_i, \dot{w}_i' \rangle = 0,\ \text{where } \dot{w}_i' \in V_{w_i}\big\}. \end{aligned} \tag{A11}$$

Note that the operation $\langle \cdot, \cdot \rangle$ is not the Hermitian form defined in Eq. (8). It is instead the standard inner product in the space of complex matrices seen as a real vector space $\langle A, B \rangle \equiv \mathrm{Re\,Tr}(A^\dagger B)$. The condition in (A11) is then given by

$$\begin{aligned} \mathrm{Re\,Tr}\left(\dot{w}_i^\dagger w_i \cdot X\right) = 0 \quad &\text{for every } X \in \mathfrak{u}(r_i) \\ \Rightarrow \dot{w}_i^\dagger w_i - w_i^\dagger \dot{w}_i = 0, \end{aligned} \tag{A12}$$

where the implication stems from the fact that X is anti-Hermitian, so that $\dot{w}^\dagger w$ can only be Hermitian [36]. We can go further by making use of the condition in Eq. (A5), yielding $\dot{w}_i^\dagger w_i = -w_i^\dagger \dot{w}_i$; substituting this into Eq. (A12), we get

$$\dot{w}_i^\dagger w_i - w_i^\dagger \dot{w}_i = -2 w_i^\dagger \dot{w}_i = 0 \Rightarrow w_i^\dagger \dot{w}_i = 0. \tag{A13}$$

Finally, now that we have a notion of horizontal subspaces of the tangent spaces to $V_{r_i}(\mathbb{C}^n)$, we have unique isomorphisms of $H_{w_i} \cong T_{P_i}\mathrm{Gr}_{r_i}(\mathbb{C}^n)$ provided by the projection π_i. This means that for each $v \in T_{P_i}\mathrm{Gr}_{r_i}(\mathbb{C}^n)$ there exists a unique $\widetilde{v}^H \in H_{w_i} \subset T_{w_i} \mathrm{V}_{r_i}(\mathbb{C}^n)$, such that its projection is v, i.e., $\pi_i(\widetilde{v}^H) = \widetilde{v}^H w_i^\dagger + w_i \widetilde{v}^{H\dagger} = v$, and the converse is also true. This lift is called the "horizontal lift" for obvious reasons. Any other lift of v to $T_{w_i}\mathrm{V}_{r_i}(\mathbb{C}^n)$, i.e., any tangent vector projecting to v, would differ from the horizontal by an element of the kernel of the derivative of the projection, i.e., a vertical vector. As a consequence of this isomorphism, the Riemannian metric in the base space is $g_i(v_1, v_2) := \langle \widetilde{v}_1^H, \widetilde{v}_2^H \rangle = \mathrm{Re\,Tr}\,[(\widetilde{v}_1^H)^\dagger \widetilde{v}_2^H]$, where $\widetilde{v}_i^H$ are horizontal lifts of tangent vectors $v_1, v_2 \in T_{P_i}\mathrm{Gr}_{r_i}(\mathbb{C}^n)$. Moreover, the expression $g_i(v_1, v_2)$ does not depend on the point of the fiber over P_i, because the horizontal subspaces are $\mathrm{U}(r_i)$ equivariant and the metric is $\mathrm{U}(r_i)$ invariant. Indeed, if $\widetilde{v}^H \in H_{w_i}$ is a horizontal lift of $v \in T_{P_i}\mathrm{Gr}_{r_i}(\mathbb{C}^n)$, then $\widetilde{v}^H \cdot U$ is a horizontal lift belonging to $H_{w_i \cdot U}$, for every $U \in \mathrm{U}(r_i)$: $w_i^\dagger \widetilde{v}^H = 0 \Rightarrow (w_i \cdot U)^\dagger (\widetilde{v}^H \cdot U) = U^\dagger w_i^\dagger \widetilde{v}^H U = 0$. Note that in $\widetilde{v}^H \cdot U$, right multiplication should be understood as the tangent map of right multiplication at w_i. Finally, $\mathrm{Re\,Tr}\,[(\widetilde{v}_1^H)^\dagger \widetilde{v}_2^H] = \mathrm{Re\,Tr}\,[(\widetilde{v}_1^H \cdot U)^\dagger \widetilde{v}_2^H \cdot U]$, by the cyclic property of the trace, which shows that this expression defines a metric in the base space.

Now every tangent vector $\widetilde{v} \in T_{w_i}\mathrm{V}_{r_i}(\mathbb{C}^n)$ is uniquely projected to a horizontal vector $\widetilde{v}^H \in H_{w_i}$, which is mapped to a base space tangent vector $v \in T_{P_i}\mathrm{Gr}_{r_i}(\mathbb{C}^n)$. Given the decomposition $T_{w_i}\mathrm{V}_{r_i}(\mathbb{C}^n) = V_{w_i} \oplus H_{w_i}$, we can always find unique projection operators onto the vertical and horizontal subspaces that perform the splitting

$$\widetilde{v} = \widetilde{v}^V + \widetilde{v}^H \quad \text{where } \widetilde{v}^V \in V_{w_i}, \widetilde{v}^H \in H_{w_i}. \tag{A14}$$

We have the identity

$$g(v_1, v_2) = \langle \widetilde{v}_1^H, \widetilde{v}_2^H \rangle. \tag{A15}$$

Additionally, due to the splitting of subspaces, we can write

$$\widetilde{v}^H = \widetilde{v} - \widetilde{v}^V. \tag{A16}$$

In the following, we determine the form of the projection onto the vertical subspaces, in order to obtain a more compact form for the metric on the base space.

We claim that the vertical projection of a general tangent vector $\widetilde{v}$ is given by

$$\widetilde{v}^V = P_i \widetilde{v} = w_i w_i^\dagger \widetilde{v}. \tag{A17}$$

Let us see why this is true. For this tangent vector to be vertical it must comply with Eq. (A7), i.e.,

$$(P_i\widetilde{v})w_i^\dagger + w_i(P_i\widetilde{v})^\dagger = w_i w_i^\dagger \widetilde{v} w_i^\dagger + w_i \widetilde{v}^\dagger w_i w_i^\dagger = 0. \quad \text{(A18)}$$

However, we know that $\widetilde{v}$ is a tangent vector; that is, we know that $\widetilde{v}^\dagger w_i = -w_i^\dagger \widetilde{v}$. Replacing this in the expression above, we have

$$w_i w_i^\dagger \widetilde{v} w_i^\dagger - w_i w_i^\dagger \widetilde{v} w_i^\dagger = 0. \quad \text{(A19)}$$

Hence we have verified that $P_i\widetilde{v}$ is a vertical tangent vector and the map $\widetilde{v} \mapsto w_i w_i^\dagger \widetilde{v}$ is a projection onto the vertical space. The horizontal projection is then given by

$$\widetilde{v}^H = \widetilde{v} - (w_i w_i^\dagger)\widetilde{v}. \quad \text{(A20)}$$

Meanwhile, the metric in $\mathrm{Gr}_{r_i}(\mathbb{C}^n)$ is, using the horizontal projections, given by the following compact formula:

$$\begin{aligned} g_i &= \mathrm{Re\,Tr}[(dw_i^\dagger - dw_i^\dagger w_i w_i^\dagger)(dw_i - w_i w_i^\dagger dw_i)] \\ &= \mathrm{Re\,Tr}[dw_i^\dagger dw_i - dw_i^\dagger w_i w_i^\dagger dw_i \\ &\quad - dw_i^\dagger w_i w_i^\dagger dw_i + dw_i^\dagger (w_i w_i^\dagger)^2 dw_i]. \end{aligned} \quad \text{(A21)}$$

We know that $w_i(w_i^\dagger w_i)w_i^\dagger = w_i w_i^\dagger$, since $w_i^\dagger w_i = I_k$, so the last two terms cancel each other, giving

$$\begin{aligned} g_i &= \mathrm{Re\,Tr}[dw_i^\dagger dw_i - dw_i^\dagger w_i w_i^\dagger dw_i] \\ &= \mathrm{Re\,Tr}[dw_i^\dagger (1 - w_i w_i^\dagger) dw_i]. \end{aligned} \quad \text{(A22)}$$

Now, this expression is written in terms of the elements defined in the principal bundle, so we want to write it in terms of the elements in the base space—the projectors P_i. For this purpose, notice that $w_i = (w_i w_i^\dagger)w_i = P_i w_i$, which by derivation gives $dw_i = dP_i w_i + P_i dw_i$. The same can be done for the Hermitian $w_i^\dagger = w_i^\dagger (w_i w_i^\dagger) = w_i^\dagger P_i$, which gives us $dw_i^\dagger = dw_i^\dagger P_i + w_i^\dagger dP_i$. Replacing these in Eq. (A22), we get

$$\begin{aligned} g_i &= \mathrm{Re\,Tr}[dw_i^\dagger (1 - w_i w_i^\dagger) dw_i] \\ &= \mathrm{Re\,Tr}[(dw_i^\dagger P_i + w_i^\dagger dP_i)(1 - P_i)(dP_i w_i + P_i dw_i)] \\ &= \mathrm{Re\,Tr}[(dw_i^\dagger P_i + w_i^\dagger dP_i - dw_i^\dagger P_i - w_i^\dagger dP_i P_i) \\ &\quad \cdot (dP_i w_i + P_i dw_i)] \\ &= \mathrm{Re\,Tr}(dw_i^\dagger P_i dP_i w_i + w_i^\dagger dP_i dP_i w_i \\ &\quad - dw_i^\dagger P_i dP_i w_i - w_i^\dagger dP_i P_i dP_i w_i \\ &\quad + dw_i^\dagger P_i dw_i + w_i^\dagger dP_i P_i dw_i \\ &\quad - dw_i^\dagger P_i dw_i - w_i^\dagger dP_i P_i dw_i) \\ &= \mathrm{Re\,Tr}(w_i^\dagger dP_i dP_i w_i - w_i^\dagger dP_i P_i dP_i w_i) \\ &= \mathrm{Re\,Tr}(P_i dP_i dP_i) - \mathrm{Re\,Tr}(P_i dP_i P_i dP_i). \end{aligned} \quad \text{(A23)}$$

Moreover, since $P_i^2 = P_i$, we have that $dP_i = d(P_i^2) = P_i dP_i + dP_i P_i$. Multiplying this expression by P_i on both sides, we get $P_i dP_i P_i = 2P_i dP_i P_i$, and we can conclude that $P_i dP_i P_i = 0$. The last term on the last expression is then zero, and we see that the metric is given by

$$\begin{aligned} g_i &= \mathrm{Re\,Tr}\,(P_i dP_i dP_i) = \mathrm{Re\,Tr}\,(P_i dP_i dP_i P_i) \\ &= \mathrm{Tr}\,(P_i dP_i dP_i P_i). \end{aligned} \quad \text{(A24)}$$

Now we wish to determine the metric on the total space of the principal bundle, i.e., the metric that encompasses both the classical and quantum parts. For this purpose, consider a curve in the principal bundle space given by $t \mapsto p_\tau(t) = (\sqrt{\mathbf{p}(t)}, \mathbf{w}(t))$ and compute the distance between two infinitesimally close points t and $t + \delta t$. For the first case, we consider a static $\mathbf{w}(t) = \mathbf{w}$ and compute the distance

$$\begin{aligned} &d_\tau^2(p_\tau(t),\ p_\tau(t+\delta t)) \\ &\quad = 2\left(1 - \sum_{i=1}^k \sqrt{p_i(t)p_i(t+\delta t)}\,\mathrm{Re\,Tr}(w_i^\dagger w_i)\right). \end{aligned} \quad \text{(A25)}$$

We have $\mathrm{Tr}(w_i^\dagger w_i) = \mathrm{Tr}\,P_i = r_i$; hence

$$d_\tau^2(p_\tau(t), p_\tau(t+\delta t)) = 2\left(1 - \sum_{i=1}^k r_i \sqrt{p_i(t)p_i(t+\delta t)}\right). \quad \text{(A26)}$$

Let us look more closely at the expression $\sqrt{p_i(t)p_i(t+\delta t)}$. We can Taylor expand $P_i(t+\delta t)$ to second order in δt to obtain

$$\begin{aligned} \sqrt{p_i(t)p_i(t+\delta t)} &= \sqrt{p_i(t)\left(p_i(t) + \dot{p}_i \delta t + \frac{1}{2}\ddot{p}_i \delta t^2\right)} \\ &= p_i(t)\sqrt{1 + \frac{\dot{p}_i}{p_i}\delta t + \frac{1}{2}\frac{\ddot{p}_i}{p_i}\delta t^2}. \end{aligned} \quad \text{(A27)}$$

We can then approximate the quantity inside the square root by $\sqrt{1+x} \approx 1 + \frac{1}{2}x - \frac{1}{8}x^2$, which, ignoring higher-order terms, yields

$$\begin{aligned} &\sqrt{p_i(t)p_i(t+\delta t)} \\ &\quad \approx p_i\left[1 + \frac{1}{2}\left(\frac{\dot{p}_i}{p_i}\delta t + \frac{1}{2}\frac{\ddot{p}_i}{p_i}\delta t^2\right) - \frac{1}{8}\left(\frac{\dot{p}_i}{p_i}\delta t + \frac{1}{2}\frac{\ddot{p}_i}{p_i}\delta t^2\right)^2\right] \\ &\quad = p_i\left[1 + \frac{1}{2}\frac{\dot{p}_i}{p_i}\delta t + \frac{1}{2}\frac{\ddot{p}_i}{p_i}\delta t^2 - \frac{1}{8}\left(\frac{\dot{p}_i}{p_i}\right)^2 \delta t^2\right] \\ &\quad = p_i + \frac{1}{2}\dot{p}_i \delta t + \frac{1}{2}\ddot{p}_i \delta t^2 - \frac{1}{8}\frac{\dot{p}_i^2}{p_i}\delta t^2. \end{aligned} \quad \text{(A28)}$$

Replacing this in Eq. (A26), we get

$$\begin{aligned} &d_\tau^2(p_\tau(t), p_\tau(t+\delta t)) \\ &\quad = 2\left[1 - \sum_{i=1}^k r_i\left(p_i + \frac{1}{2}\dot{p}_i \delta t + \frac{1}{2}\ddot{p}_i \delta t^2 - \frac{1}{8}\frac{\dot{p}_i^2}{p_i}\delta t^2\right)\right]. \end{aligned} \quad \text{(A29)}$$

Using the condition $\sum_{i=1}^k r_i p_i = 1$, we can infer that $\sum_{i=1}^k r_i \dot{p}_i = 0$ and $\sum_{i=1}^k r_i \ddot{p}_i = 0$. Applying these results in the expression above, we finally arrive at the Fisher-Rao metric

$$\left(ds_P^{\mathrm{Cl}}\right)^2 = \frac{1}{4}\sum_{i=1}^k r_i \frac{\dot{p}_i^2}{p_i}\delta t^2 = \frac{1}{4}\sum_{i=1}^k r_i \frac{dp_i^2}{p_i} \quad \text{(A30)}$$

in terms of the probability distribution "coordinates" $\sqrt{\mathbf{p}}$.

Next, consider the case of a static classical part $\mathbf{p}(t) = \mathbf{p}$. The distance is then

$$d_\tau^2(p_\tau(t),\ p_\tau(t+\delta t)) = 2\left(1 - \sum_{i=1}^{k} p_i \,\mathrm{Re\,Tr}[w_i(t)^\dagger w_i(t+\delta t)]\right). \tag{A31}$$

Expanding $w_i(t+\delta t)$ to second order, $w_i(t+\delta t) \approx w_i(t) + \dot{w}_i(t)\delta t + \frac{1}{2}\ddot{w}_i(t)\delta t^2$, we have

$$\begin{aligned}&\mathrm{Re\,Tr}[w_i(t)^\dagger w_i(t+\delta t)] \\ &= \mathrm{Re\,Tr}(w_i^\dagger w_i) + \mathrm{Re\,Tr}(w_i^\dagger \dot{w}_i)\delta t + \tfrac{1}{2}\,\mathrm{Re\,Tr}(w_i^\dagger \ddot{w}_i)\delta t^2 \\ &\quad + r_i + \tfrac{1}{2}\,\mathrm{Tr}(w_i^\dagger \dot{w}_i + \dot{w}_i^\dagger w_i)\delta t + \tfrac{1}{4}\,\mathrm{Tr}(w_i^\dagger \ddot{w}_i + \ddot{w}_i^\dagger w_i)\delta t^2.\end{aligned} \tag{A32}$$

From condition (A5) for tangent vectors, the first-order term is zero. From this same condition, one can infer that $\ddot{w}_i^\dagger w_i + w_i^\dagger \ddot{w}_i = -2\dot{w}_i^\dagger \dot{w}_i$, and Eq. (A32) becomes

$$\mathrm{Re\,Tr}[w_i(t)^\dagger w_i(t+\delta t)] = r_i - \tfrac{1}{2}\,\mathrm{Tr}(\dot{w}_i^\dagger \dot{w}_i)\delta t^2. \tag{A33}$$

Using this expression in Eq. (A31), we get

$$d_\tau^2(p_\tau(t),\ p_\tau(t+\delta t)) = 2\left(1 - \sum_{i=1}^{k} r_i p_i + \frac{1}{2}\sum_{i=1}^{k} p_i \,\mathrm{Tr}(\dot{w}_i^\dagger \dot{w}_i)\delta t^2\right). \tag{A34}$$

Since $\sum_{i=1}^{k} r_i p_i = 1$, we have

$$\left(ds^{\mathrm{Q}}_{P_\tau}\right)^2 = \sum_{i=1}^{k} p_i \,\mathrm{Tr}(\dot{w}_i^\dagger \dot{w}_i)\delta t^2 = \sum_{i=1}^{k} p_i \,\mathrm{Tr}(dw_i^\dagger dw_i). \tag{A35}$$

From the derivation of Eq. (A24), it becomes clear that, restricting to the horizontal subspaces, one obtains the induced quantum part of the metric in the base space

$$\left(ds^{\mathrm{Q}}_{B_\tau}\right)^2 = \sum_{i=1}^{k} p_i \,\mathrm{Tr}\,(P_i dP_i dP_i). \tag{A36}$$

So, the quantum part of the metric in the base space is the sum for $i \in \{1,\ldots,k\}$ of the metric on the Grassmannian given by Eq. (A24) weighed by the relative proportions of the distribution p_i.

Finally, we are left with the task of taking a general variation, where both $\sqrt{\mathbf{p}}(t)$ and $\mathbf{w}(t)$ are nonconstant, to make sure that we do not get cross terms. We have

$$d_\tau^2(p_\tau(t), p_\tau(t+\delta t)) = 2\left(1 - \sum_{i=1}^{k} \sqrt{p_i(t)p_i(t+\delta t)}\,\mathrm{Re\,Tr}[w_i(t)^\dagger w_i(t+\delta t)]\right). \tag{A37}$$

We can Taylor expand, as before, to obtain

$$\begin{aligned}&d_\tau^2(p_\tau(t),\ p_\tau(t+\delta t)) \\ &= 2\left\{1 - \sum_{i=1}^{k}\left[p_i + \frac{1}{2}\dot{p}_i\delta t + \frac{1}{2}\ddot{p}_i\delta t^2\right.\right. \\ &\quad \left.\left. - \frac{1}{8}\frac{\dot{p}_i^2}{p_i}\delta t^2\left(r_i - \frac{1}{2}\,\mathrm{Tr}(\dot{w}_i^\dagger \dot{w}_i)\delta t^2\right)\right]\right\}.\end{aligned} \tag{A38}$$

Collecting the terms up to second order, we get

$$\begin{aligned}&d_\tau^2(p_\tau(t), p_\tau(t+\delta t)) \\ &= 2\left[1 - \sum_{i=1}^{k}\left(p_i \,\mathrm{Tr}(\dot{w}_i^\dagger \dot{w}_i)\delta t^2 + \frac{1}{2}r_i\dot{p}_i\delta t\right.\right. \\ &\quad \left.\left. + \frac{1}{2}r_i\ddot{p}_i\delta t^2 - \frac{1}{8}r_i\frac{\dot{p}_i^2}{p_i}\delta t^2\right)\right],\end{aligned} \tag{A39}$$

which, using the same arguments as before, reduces to

$$\begin{aligned}ds^2_{P_\tau} &= \sum_{i=1}^{k}\left(\frac{1}{4}r_i\frac{\dot{p}_i^2}{p_i}\delta t^2 + p_i \,\mathrm{Tr}(\dot{w}_i^\dagger \dot{w}_i)\delta t^2\right) \\ &= \sum_{i=1}^{k}\left(\frac{1}{4}r_i\frac{dp_i^2}{p_i} + p_i \,\mathrm{Tr}(dw_i^\dagger dw_i)\right).\end{aligned} \tag{A40}$$

Hence the metric in the principal bundle is just the sum of the respective classical and quantum metrics. We want to arrive at the metric for the base space: The classical probability distributions $\sqrt{p_i}$ have no gauge freedom, so they have no vertical or horizontal components and their projection is trivial; meanwhile, the horizontal projection in the quantum part described by the amplitudes w_i proceeds as in the Stiefel manifold case, for each $i = 1,\ldots,k$, so that our final interferometric metric g_I is

$$\begin{aligned}g_I &= ds^2_{B_\tau} \\ &= \left(ds^{\mathrm{Cl}}_{B_\tau}\right)^2 + \left(ds^{\mathrm{Q}}_{B_\tau}\right)^2 \\ &= \frac{1}{4}\sum_{i=1}^{k} r_i\frac{dp_i^2}{p_i} + \sum_{i=1}^{k} p_i \,\mathrm{Tr}\,(P_i dP_i dP_i).\end{aligned} \tag{A41}$$

APPENDIX B: THE PROOF OF THE MAXIMAL OUTPUT PROBABILITY IN THE INTERFEROMETRIC EXPERIMENT

The input state is $|0\rangle\langle 0| \otimes \rho$. The first beam splitter BS1 $\otimes$ I acts on this state giving $\frac{1}{2}(|0\rangle + i|1\rangle)(\langle 0| - i\langle 1|) \otimes \rho$. The controlled unitary is $|0\rangle\langle 0| \otimes V + |1\rangle\langle 1| \otimes U$, which, when acting on the last state, gives

$$\begin{aligned}&\tfrac{1}{2}(|0\rangle\langle 0| \otimes V\rho V^\dagger - i|0\rangle\langle 1| \otimes V\rho U^\dagger \\ &\quad + i|1\rangle\langle 0| \otimes U\rho V^\dagger + |1\rangle\langle 1| \otimes U\rho U^\dagger).\end{aligned} \tag{B1}$$

Passing through a second beam splitter and measuring the $|1\rangle$ state yield

$$\tfrac{1}{4}|1\rangle\langle 1| \otimes [V\rho V^\dagger + V\rho U^\dagger + U\rho V^\dagger + U\rho U^\dagger]. \tag{B2}$$

Tracing out this quantity gives

$$\tfrac{1}{4}[\mathrm{Tr}\,U\rho U^\dagger + \mathrm{Tr}\,V\rho V^\dagger + 2\,\mathrm{Re\,Tr}\,U\rho V^\dagger]. \tag{B3}$$

We know that $\mathrm{Tr}\,U\rho U^\dagger = \mathrm{Tr}\,V\rho V^\dagger = 1$; hence

$$\tfrac{1}{2}[1 + \mathrm{Re\,Tr}\,U\rho V^\dagger]. \tag{B4}$$

Recall that $V = \sum_{i=0}^{k} P_i V P_i$ and that since we can write, in terms of a choice of amplitudes w_i, $i = 1, \ldots, k$,

$$P_i = w_i w_i^\dagger, \quad i = 1, \ldots, k, \tag{B5}$$

then

$$V = P_0 V P_0 + \sum_{i=1}^{k} w_i V_i w_i^\dagger, \tag{B6}$$

where $V_i = w_i^\dagger V w_i$ is an $r_i \times r_i$ unitary matrix, for $i = 1, \ldots, k$. Observe that

$$\begin{aligned} \mathrm{Tr}[V^\dagger U \rho] &= \sum_{i,j=0}^{k} p_i \,\mathrm{Tr}[P_j V^\dagger P_j U P_i] \\ &= \sum_{i=0}^{k} p_i \,\mathrm{Tr}[V^\dagger P_i U P_i], \end{aligned} \tag{B7}$$

where in the last step we used the cyclic property of the trace and $P_i P_j = \delta_{ij} P_i$, $i, j = 0, \ldots, k$. Finally, introducing the expression for V of Eq. (B6), we can write, using $w_i^\dagger w_i = I_{r_i}$, $i = 1, \ldots, r_i$, and $p_0 = 0$,

$$\begin{aligned} \sum_{i=1}^{k} p_i \,\mathrm{Tr}[V^\dagger P_i U P_i] &= \sum_{i=1}^{k} p_i \,\mathrm{Tr}[(V_i^\dagger w_i^\dagger U) w_i] \\ &= \sum_{i=1}^{k} p_i \,\mathrm{Tr}[(U^\dagger w_i V_i)^\dagger w_i]. \end{aligned} \tag{B8}$$

Observe that if we write

$$p_\tau = ((p_i, w_i))_{i=1}^{k} \text{ and } q_\tau = ((p_i, U^\dagger w_i V_i))_{i=1}^{k}, \tag{B9}$$

then

$$\sum_{i=1}^{k} p_i \,\mathrm{Tr}[V_i^\dagger w_i^\dagger U w_i] = \langle q_\tau, p_\tau \rangle_\tau, \tag{B10}$$

where $\langle q_\tau, p_\tau \rangle$ is the Hermitian form defined in Eq. (8). Hence

$$\begin{aligned} \mathrm{pr}_1 &= \frac{1}{2}(1 + \mathrm{Re}\,\mathrm{Tr}\, U \rho V^\dagger) \\ &= 1 - \frac{1}{2}\left(1 - \sum_{i=1}^{k} p_i \,\mathrm{Re}\,\mathrm{Tr}\left[P_i V^\dagger P_i U P_i\right]\right) \\ &= 1 - \frac{1}{2}\left(1 - \sum_{i=1}^{k} p_i \,\mathrm{Re}\langle q_\tau, p_\tau \rangle_\tau\right) \\ &= 1 - \frac{1}{4} d_\tau^2(q_\tau, p_\tau), \end{aligned} \tag{B11}$$

where d_τ is the distance over the total space of the principal bundle $P_\tau \to B_\tau$. Maximizing over the the gauge degree of freedom given by the collection of unitary $r_i \times r_i$ matrices, V_i, $i = 1, \ldots, k$ (note that $P_0 V P_0$ is irrelevant), one gets the distance $d_I(\rho, U^\dagger \rho U)$ over the base space B_τ as explored in the main text.

APPENDIX C: PULLBACK OF INTERFEROMETRIC TO PARAMETER SPACE

We wish to find the metric obtained by pulling back the interferometric metric

$$\begin{aligned} g &= \frac{1}{4} \sum_i r_i \frac{dp_i^2}{p_i} + \sum_i p_i \,\mathrm{Tr}\,(P_i dP_i dP_i) \\ &\text{with } \rho = \sum_i p_i P_i \text{ and } \mathrm{Tr}\, P_i = r_i, \end{aligned} \tag{C1}$$

by the map induced by the Gibbs state

$$M \mapsto \rho(M) = Z^{-1} \exp[-\beta \mathcal{H}(M)], \tag{C2}$$

with $\mathcal{H}(M)$ given by Eq. (19) and where Z is the partition function. The first thing to notice is that if $\rho = \rho_1 \otimes \rho_2$, with $\rho_\alpha = \sum_{i_\alpha} p_{i_\alpha} P_{i_\alpha}$, $\alpha = 1, 2$ we have the decomposition

$$\rho = \sum_I p_I P_I = \sum_{i_1, i_2} p_{i_1} p_{i_2} P_{i_1} \otimes P_{i_2}, \tag{C3}$$

where $I = (i_1, i_2)$ is a multi-index describing the joint system labels. Note that

$$\begin{aligned} &\sum_{i_1, i_2} r_I \frac{dp_I^2}{p_I} \\ &\quad = \sum_{i_1, i_2} \frac{r_{i_1} r_{i_2}}{p_{i_1} p_{i_2}} \left(p_{i_2}^2 dp_{i_1} dp_{i_1} + 2 p_{i_1} p_{i_2} dp_{i_1} dp_{i_2} + p_{i_1}^2 dp_{i_2}^2\right) \\ &\quad = \sum_{i_1} r_{i_1} \frac{dp_{i_1}^2}{p_{i_1}} + \sum_{i_2} r_{i_2} \frac{dp_{i_2}^2}{p_{i_2}} \end{aligned} \tag{C4}$$

and

$$\begin{aligned} &\sum_I p_I \,\mathrm{Tr}\,(P_I dP_I dP_I) \\ &\quad = \sum_{i_1, i_2} p_{i_1} p_{i_2} \,\mathrm{Tr}\left[P_{i_1} \otimes P_{i_2} d\left(P_{i_1} \otimes P_{i_2}\right) d\left(P_{i_1} \otimes P_{i_2}\right)\right] \\ &\quad = \sum_{i_1} p_{i_1} \,\mathrm{Tr}\left(P_{i_1} dP_{i_1} dP_{i_1}\right) + \sum_{i_2} p_{i_2} \,\mathrm{Tr}\left(P_{i_2} dP_{i_2} dP_{i_2}\right), \end{aligned} \tag{C5}$$

where we used $PdPP = 0$ for any projector P. As a consequence, the interferometric metric, much like the Bures metric, converts tensor product states into orthogonal sum metrics.

Because the Hamiltonian is diagonal in momentum space, the density matrix factors over the momenta: It follows that the metric becomes an integral over the momentum space of individual contributions of each momentum sector. The pullback of the classical term, which also appears in the Bures metric,

$$\frac{1}{4} \sum_i r_i \frac{dp_i^2}{p_i}, \tag{C6}$$

was computed in the Appendix of Ref. [22], and it yields

$$\frac{\beta^2}{4} \int_{\mathrm{BZ}^d} \frac{d^d k}{(2\pi)^d} \frac{1}{\cosh[\beta E(\mathbf{k}; M)] + 1} \left(\frac{\partial E(\mathbf{k}; M)}{\partial M}\right)^2 dM^2, \tag{C7}$$

where $E(\mathbf{k};M) = |d(\mathbf{k};M)|$ is the magnitude of $d(\mathbf{k},M)$. With regard to the second term, one can use the mathematical fact that the embedding of the space of k-dimensional subspaces of $\mathbb{C}^N$, $\mathrm{Gr}_k(\mathbb{C}^N)$ on the space of one-dimensional subspaces of the Fock space $\mathbb{P}\Lambda^*\mathbb{C}^N$, given by

$$\mathrm{span}\{|1\rangle,\ldots,|k\rangle\} \mapsto \mathrm{span}\{c_1^\dagger \cdots c_k^\dagger|0\rangle\}, \tag{C8}$$

is isometric. In the previous equation, $c_i^\dagger$ stand for creation operators for $|i\rangle$; that is, at the single-particle level, $c_i^\dagger|0\rangle = |i\rangle$, $i = 1,\ldots,k$. The embedding being isometric means, in this context, that if we write the rank k single-particle projector

$$\widetilde{P} = \sum_{i=1}^{k} |i\rangle\langle i| \tag{C9}$$

and the rank 1 many-body projector

$$P = c_1^\dagger \cdots c_k^\dagger|0\rangle\langle 0|c_k \cdots c_1, \tag{C10}$$

we have

$$\mathrm{Tr}(\widetilde{P}d\widetilde{P}d\widetilde{P}) = \mathrm{Tr}(PdPdP). \tag{C11}$$

In particular, this means that in the gapped case for each $\mathbf{k} \in \mathrm{BZ}^d$ we will have four classes of orthogonal eigenstates,

$$|0\rangle, c_{1,\mathbf{k}}^\dagger|0\rangle, c_{2,\mathbf{k}}^\dagger|0\rangle, c_{1,\mathbf{k}}^\dagger c_{2,\mathbf{k}}^\dagger|0\rangle, \tag{C12}$$

where $c_{i,\mathbf{k}}^\dagger$, $i = 1,2$, are the Bogoliubov quasiparticle creation operators of $\mathcal{H}$ with energies $E(\mathbf{k};M)$ and $-E(\mathbf{k};M)$, respectively. The energies of the classes of eigenstates are 0, $E(\mathbf{k};M)$, $-E(\mathbf{k};M)$, and 0, respectively. The associated single-particle 2×2 projectors are the 0 projector, $P_1(\mathbf{k};M) = c_{1,\mathbf{k}}^\dagger|0\rangle\langle 0|c_{1,\mathbf{k}}$, $P_2(\mathbf{k};M) = c_{2,\mathbf{k}}^\dagger|0\rangle\langle 0|c_{2,\mathbf{k}}$, and the 2×2 identity matrix I_2, respectively. Only $P_1(\mathbf{k})$ and $P_2(\mathbf{k})$ are nontrivial (not constant), and moreover, if we introduce the unit vector $n = d/|d|$, we can write

$$\begin{aligned} P_1(\mathbf{k};M) &= \tfrac{1}{2}[I_2 + n^\mu(\mathbf{k};M)\sigma_\mu], \\ P_2(\mathbf{k};M) &= I_2 - P_1(\mathbf{k};M). \end{aligned} \tag{C13}$$

As a consequence, using the identity $\mathrm{Tr}\,(PdPdP) = (1/2)\,\mathrm{Tr}\,(dPdP)$ and using the fact that the Pauli matrices are traceless, we get

$$\begin{aligned} \mathrm{Tr}\,(P_1dP_1dP_1) &= \mathrm{Tr}(P_2dP_2dP_2) \\ &= \frac{1}{4}\delta_{\mu\nu}\frac{\partial n^\mu(\mathbf{k};M)}{\partial M}\frac{\partial n^\nu(\mathbf{k};M)}{\partial M}dM^2. \end{aligned} \tag{C14}$$

Finally, taking into account the partition function factor $Z_\mathbf{k} = \{2 + 2\cosh[\beta E(\mathbf{k};M)]\}$, we get that the quantum contribution is

$$\begin{aligned} &\frac{1}{4}\int_{\mathrm{BZ}^d}\frac{d^dk}{(2\pi)^d}\left(\frac{\cosh[\beta E(\mathbf{k};M)]}{1+\cosh[\beta E(\mathbf{k};M)]}\right) \\ &\quad\times\delta_{\mu\nu}\frac{\partial n^\mu(\mathbf{k};M)}{\partial M}\frac{\partial n^\nu(\mathbf{k};M)}{\partial M}dM^2. \end{aligned}$$

Finally, we obtain

$$\begin{aligned} g = \frac{1}{4}\int_{\mathrm{BZ}^d}\frac{d^dk}{(2\pi)^d}\Bigg\{&\frac{1}{\cosh(\beta E)+1} \\ &\times\left[\beta^2\left(\frac{\partial E}{\partial M}\right)^2 + \cosh(\beta E)\delta_{\mu\nu}\frac{\partial n^\mu}{\partial M}\frac{\partial n^\nu}{\partial M}\right]\Bigg\}dM^2, \end{aligned} \tag{C15}$$

where we omitted the obvious dependence on $\mathbf{k}$ and M of the quantities E and n^μ.

[1] F. D. M. Haldane, Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly", Phys. Rev. Lett. **61**, 2015 (1988).

[2] A. Kitaev, Periodic table for topological insulators and superconductors, in *Advances in Theoretical Physics: Landau Memorial Conference*, AIP Conference Proceedings Vol. 1134 (American Institute of Physics, Melville, NY, 2009), pp. 22–30.

[3] K. Shiozaki and M. Sato, Topology of crystalline insulators and superconductors, Phys. Rev. B **90**, 165114 (2014).

[4] Z. Gong, Y. Ashida, K. Kawabata, K. Takasan, S. Higashikawa, and M. Ueda, Topological Phases of Non-Hermitian Systems, Phys. Rev. X **8**, 031079 (2018).

[5] R. Roy and F. Harper, Periodic table for Floquet topological insulators, Phys. Rev. B **96**, 155118 (2017).

[6] F. Schindler, A. M. Cook, M. G. Vergniory, Z. Wang, S. S. P. Parkin, B. A. Bernevig, and T. Neupert, Higher-order topological insulators, Sci. Adv. **4**, eaat0346 (2018).

[7] L. D. Landau, On the theory of phase transitions. i., Zh. Eksp. Teor. Fiz. **7**, 19 (1937).

[8] P. Zanardi and N. Paunković, Ground state overlap and quantum phase transitions, Phys. Rev. E **74**, 031123 (2006).

[9] N. Paunković, P. D. Sacramento, P. Nogueira, V. R. Vieira, and V. K. Dugaev, Fidelity between partial states as a signature of quantum phase transitions, Phys. Rev. A **77**, 052302 (2008).

[10] N. Paunković and V. R. Vieira, Macroscopic distinguishability between quantum states defining different phases of matter: Fidelity and the Uhlmann geometric phase, Phys. Rev. E **77**, 011129 (2008).

[11] P. Zanardi, L. Campos Venuti, and P. Giorda, Bures metric over thermal state manifolds and quantum criticality, Phys. Rev. A **76**, 062318 (2007).

[12] P. Zanardi, P. Giorda, and M. Cozzini, Information-Theoretic Differential Geometry of Quantum Phase Transitions, Phys. Rev. Lett. **99**, 100603 (2007).

[13] L. Campos Venuti and P. Zanardi, Quantum Critical Scaling of the Geometric Tensors, Phys. Rev. Lett. **99**, 095701 (2007).

[14] W.-L. You, Y.-W. Li, and S.-J. Gu, Fidelity, dynamic structure factor, and susceptibility in critical phenomena, Phys. Rev. E **76**, 022101 (2007).

[15] A. Carollo, B. Spagnolo, and D. Valenti, Uhlmann curvature in dissipative phase transitions, Sci. Rep. **8**, 9852 (2018).

[16] T. Ozawa and N. Goldman, Extracting the quantum metric tensor through periodic driving, Phys. Rev. B **97**, 201117(R) (2018).

[17] L. Leonforte, D. Valenti, B. Spagnolo, and A. Carollo, Uhlmann number in translational invariant systems, Sci. Rep. **9**, 9106 (2019).

[18] A. Carollo, D. Valenti, and B. Spagnolo, Geometry of quantum phase transitions, Phys. Rep. **838**, 1 (2020).

[19] A. Y. Kitaev, Unpaired Majorana fermions in quantum wires, Phys.-Usp. **44**, 131 (2001).

[20] B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Uhlmann Connection in Fermionic Systems Undergoing Phase Transitions, Phys. Rev. Lett. **119**, 015702 (2017).

[21] B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Boltzmann–Gibbs states in topological quantum walks and associated many-body systems: fidelity and Uhlmann parallel transport analysis of phase transitions, J. Phys. A: Math. Theor. **50**, 365302 (2017).

[22] S. T. Amin, B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Fidelity and Uhlmann connection analysis of topological phase transitions in two dimensions, Phys. Rev. B **98**, 245141 (2018).

[23] P. Sacramento, B. Mera, and N. Paunković, Vanishing k-space fidelity and phase diagram's bulk–edge–bulk correspondence, Ann. Phys. (Amsterdam) **401**, 40 (2019).

[24] S. T. Amin, B. Mera, N. Paunković, and V. R. Vieira, Information geometric analysis of long range topological superconductors, J. Phys.: Condens. Matter **31**, 485402 (2019).

[25] U. Bhattacharya, S. Bandyopadhyay, and A. Dutta, Mixed state dynamical quantum phase transitions, Phys. Rev. B **96**, 180303(R) (2017).

[26] B. Mera, C. Vlachou, N. Paunković, V. R. Vieira, and O. Viyuela, Dynamical phase transitions at finite temperature from fidelity and interferometric Loschmidt echo induced metrics, Phys. Rev. B **97**, 094110 (2018).

[27] S. Bandyopadhyay and A. Dutta, Unitary preparation of many-body Chern insulators: Adiabatic bulk-boundary correspondence, Phys. Rev. B **102**, 094301 (2020).

[28] S. Bandyopadhyay and A. Dutta, Dissipative preparation of many-body Floquet Chern insulators, Phys. Rev. B **102**, 184302 (2020).

[29] E. Sjöqvist, A. K. Pati, A. Ekert, J. S. Anandan, M. Ericsson, D. K. L. Oi, and V. Vedral, Geometric Phases for Mixed States in Interferometry, Phys. Rev. Lett. **85**, 2845 (2000).

[30] E. Sjöqvist, Geometry along evolution of mixed quantum states, Phys. Rev. Research **2**, 013344 (2020).

[31] A. Uhlmann, Parallel transport and 'quantum holonomy' along density operators, Rep. Math. Phys. **24**, 229 (1986).

[32] S. Matsuura and S. Ryu, Momentum space metric, nonlocal operator, and topological insulators, Phys. Rev. B **82**, 245113 (2010).

[33] C.-E. Bardyn, L. Wawer, A. Altland, M. Fleischhauer, and S. Diehl, Probing the Topology of Density Matrices, Phys. Rev. X **8**, 011035 (2018).

[34] S. T. Amin, B. Mera, N. Paunković, and V. R. Vieira (unpublished).

[35] S. Morita, *Geometry of Differential Forms*, Translations of Mathematical Monographs Vol. 201 (American Mathematical Society, Providence, RI, 2001).

[36] To see this, observe that a complex matrix can be split into its Hermitian and anti-Hermitian components: $Z = Z^H + Z^{AH}$, where $Z^H = \frac{1}{2}(Z + Z^\dagger)$ and $Z^{AH} = \frac{1}{2}(Z - Z^\dagger)$. This real-linear decomposition divides the full matrix into two orthogonal components. Indeed, $\mathrm{Re}\,\mathrm{Tr}[(Z_1^{AH})^\dagger Z_2^H] = \frac{1}{2}\{\mathrm{Tr}[(Z_1^{AH})^\dagger Z_2^H] + \mathrm{Tr}[(Z_2^H)^\dagger Z_1^{AH}]\} = \frac{1}{2}\{-\mathrm{Tr}[(Z_1^{AH})Z_2^H] + \mathrm{Tr}[(Z_2^H)Z_1^{AH}]\} = 0$. Moreover, since the real vector spaces of Hermitian matrices and anti-Hermitian matrices both have dimension $k \times k$, we conclude that if a complex matrix is (real-)orthogonal to an anti-Hermitian matrix, then it must be Hermitian.

applied sciences

MDPI

Article

Generation and Distribution of Quantum Oblivious Keys for Secure Multiparty Computation

Mariano Lemus [1,2], Mariana F. Ramos [3,4], Preeti Yadav [1,2], Nuno A. Silva [3,4], Nelson J. Muga [3,4], André Souto [2,5,6,*], Nikola Paunković [1,2], Paulo Mateus [1,2] and Armando N. Pinto [3,4]

1 Departamento de Matemática, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal; mariano.lemus@tecnico.ulisboa.pt (M.L.); pri8.phy@gmail.com (P.Y.); npaunkov@math.tecnico.ulisboa.pt (N.P.); pmat@math.ist.utl.pt (P.M.)
2 Instituto de Telecomunicações, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
3 Departamento de Eletrónica, Telecomunicações e Informática, Universidade de Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal; marianaferreiraramos@live.ua.pt (M.F.R.); nasilva@ua.pt (N.A.S.); muga@ua.pt (N.J.M.); anp@ua.pt (A.N.P.)
4 Instituto de Telecomunicações, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal
5 Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa, Campo Grande 016, 1749-016 Lisboa, Portugal
6 LASIGE, Faculdade de Ciências da Universidade de Lisboa, Campo Grande 016, 1749-016 Lisboa, Portugal
* Correspondence: ansouto@fc.ul.pt

Received: 15 May 2020; Accepted: 10 June 2020; Published: 12 June 2020



Featured Application: Private data mining.

Abstract: The oblivious transfer primitive is sufficient to implement secure multiparty computation. However, secure multiparty computation based on public-key cryptography is limited by the security and efficiency of the oblivious transfer implementation. We present a method to generate and distribute oblivious keys by exchanging qubits and by performing commitments using classical hash functions. With the presented hybrid approach of quantum and classical, we obtain a practical and high-speed oblivious transfer protocol. We analyse the security and efficiency features of the technique and conclude that it presents advantages in both areas when compared to public-key based techniques.

Keywords: secure multiparty computation; oblivious transfer; quantum communications

1. Introduction

In Secure Multiparty Computation (SMC), several agents compute a function that depends on their own inputs, while maintaining them private [1]. Privacy is critical in the context of an information society, where data is collected from multiple devices (smartphones, home appliances, computers, street cameras, sensors, ...) and subjected to intensive analysis through data mining. This data collection and exploration paradigm offers great opportunities, but it also raises serious concerns. A technology able to protect the privacy of citizens, while simultaneously allowing to profit from extensive data mining, is going to be of utmost importance. SMC has the potential to be that technology if it can be made practical, secure and ubiquitous.

Current SMC protocols rely on the use of asymmetric cryptography algorithms [2], which are considered significantly more computationally complex compared with symmetric cryptography algorithms [3]. Besides being more computationally intensive, in its current standards, asymmetric cryptography cannot be considered secure anymore due to the expected increase of computational

power that a large-scale quantum computer will bring [4]. Identifying these shortcomings in efficiency and security motivates the search for alternative techniques for implementing SMC without the need of public key cryptography.

1.1. Secure Multiparty Computation and Oblivious Transfer

Consider a set of N agents and $f(x_1, x_2, \ldots, x_N) = (y_1, y_2, \ldots, y_N)$ a multivariate function. For $i \in \{1, \ldots, N\}$, a SMC service (see Figure 1) receives the input x_i from the i-th agent and outputs back the value y_i in such a way that no additional information is revealed about the remaining x_j, y_j, for $j \neq i$. Additionally, this definition can be strengthened by requiring that for some number $M < N$ of corrupt agents working together, no information about the remaining agents gets revealed (secrecy). It can also be imposed that if at most $M' < N$ agents do not compute the function correctly, the protocol identifies it and aborts (authenticity).

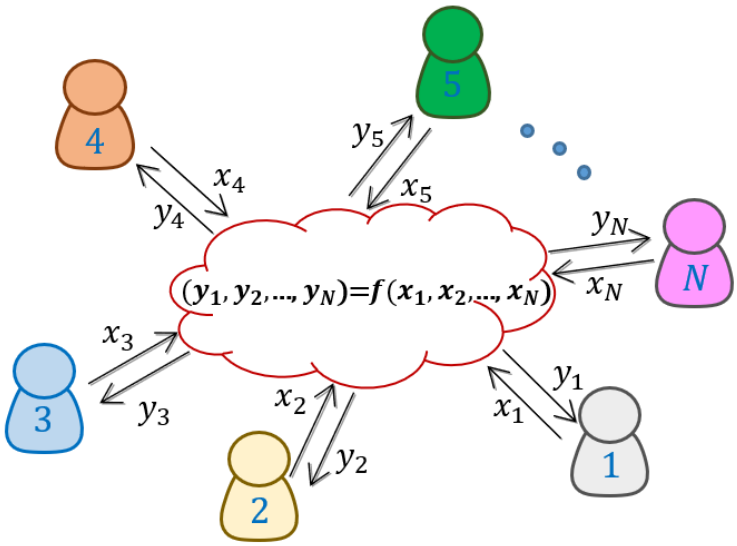


Figure 1. In secure multiparty computation, N parties compute a function preserving the privacy of their own input. Each party only has access to their own input–output pair.

Some of the most promising approaches towards implementing SMC are based on oblivious circuit evaluation techniques such as Yao's garbled circuits for the two party case [5] and the GMW or BMR protocols for the general case [2,6–8]. It has been shown that to achieve SMC it is enough to implement the Oblivious Transfer (OT) primitive and, without additional assumptions, the security of the resulting SMC depends only on that of the OT [9]. In the worst case, this requires each party to perform one OT with every other party for each gate of the circuit being evaluated. This number can be reduced by weakening the security or by increasing the amount of exchanged data [10]. Either way, the OT cost of SMC represents a major bottleneck for its practical implementation. Finding fast and secure OT protocols, hence, is a very relevant task in the context of implementing SMC.

Let Alice and Bob be two agents. A 1-out-of-2 OT service receives bits b_0, b_1 as input from Alice and a bit c as input from Bob, then outputs b_c to Bob. This is done in a way that Bob gets no information about the other message, i.e., $b_{\bar{c}}$, and Alice gets no information about Bob's choice, i.e., the value of c [11].

1.2. State of the Art

Classical OT implementations are based on the use of asymmetric keys, and suffer from two types of problems. The first one is the efficiency: asymmetric cryptography relies on relatively complex key generation, encryption, and decryption algorithms [12] (Chapter 1) and [13] (Chapter 6). This limits achievable rates of OTs, and since implementations of SMC require a very large number of OTs [3,10], this has hindered the development of SMC-based applications. The other serious drawback is that

asymmetric cryptography, based on integer number factorization or discrete-logarithm problems, is insecure in the presence of quantum computers, and therefore, it has to be progressively abandoned. There are strong research efforts in order to find other hard problems that can support asymmetric cryptography [4]. However, the security of these novel solutions is still not fully understood.

A possible way to circumvent this problem is by using quantum cryptography to improve the efficiency and security of current techniques. Quantum solutions for secure key distribution, Bit Commitment (BC) and OT have been already proposed [14]. The former was proved to be unconditionally secure (assuming an authenticated channel) and realizable using current technology. Although, it was shown to be impossible to achieve unconditionally secure quantum BC and OT [15–17], one can impose restrictions on the power of adversaries in order to obtain practically secure versions of these protocols [18,19]. These assumptions include physical limitations on the apparatuses, such as noisy or bounded quantum memories [20–22]. For instance, quantum OT and BC protocols have been developed and implemented (see [23–25]) under the noisy storage model. Nevertheless, solutions based on hardware limitations may not last for long, because as quantum technology improves the rate of secure OT instances will decrease. Other solutions include exploring relativistic scenarios using the fact that no information can travel faster than light [26–28]. However, at the moment, these solutions do not seem to be practical enough to allow the large dissemination of SMC.

In this work, we explore the resulting security and efficiency features of implementing oblivious transfer using a well known quantum protocol [5] supported by using a cryptographic hash based commitment scheme [29]. We call it a hybrid approach, since it mixes both classical and quantum cryptography. We analyse the protocol stand alone security, as well as its composable security in the random oracle model. Additionally, we study its computational complexity and compare it with the complexity of alternative public key based protocols. Furthermore, we show that, while unconditional information-theoretic security cannot be achieved, there is an advantage (both in terms of security and efficiency) of using quantum resources in computationally secure protocols, and as such, they are worth consideration for practical tasks in the near future.

This paper is organized as follows. In Section 2, we present a quantum protocol to produce OT given access to a collision resistant hash function, define the concept of oblivious keys, and explain how having pre-shared oblivious keys can significantly decrease the computational cost of OT during SMC. The security and efficiency of the protocol is discussed in Section 3. Finally, in Section 4 we summarize the main conclusions of this work.

2. Methods

2.1. Generating the OTs

In this section, we describe how to perform oblivious transfer by exchanging qubits. The protocol π_{QOT} shown in Figure 2 is the well known quantum oblivious transfer protocol first introduced by Yao, which assumes access to secure commitments. The two logical qubit states $|0\rangle$ and $|1\rangle$ represent the computational basis, and the states $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ represent the Hadamard basis. We also define the states $|(s_i, a_i)\rangle$ for $s_i, a_i \in \{0, 1\}$ according to the following rule:

$$|(0,0)\rangle = |0\rangle \quad |(0,1)\rangle = |+\rangle$$
$$|(1,0)\rangle = |1\rangle \quad |(1,1)\rangle = |-\rangle.$$

Note that these states can be physically instantiated using, for instance, a polarization encoding fiber optic quantum communication system, provided that a fast polarization encoding/decoding process and an algorithm to control random polarization drifts in optical fibers are available [30,31].

Protocol π_{QOT}

Parameters: Integers $n, m < n$.
Parties: The sender Alice and the receiver Bob.
Inputs: Alice gets two bits b_0, b_1 and Bob gets a bit c.
(Oblivious key distribution phase)

1. Alice samples $s, a \in \{0,1\}^{n+m}$. For each $i \leq n+m$ she prepares the state $|\phi_i\rangle = |(s_i, a_i)\rangle$ and sends $|\phi\rangle = |\phi_1 \phi_2 \dots \phi_{n+m}\rangle$ to Bob.
2. Bob samples $\tilde{a} \in \{0,1\}^{n+m}$ and, for each i, measures $|\phi_i\rangle$ in the computational basis if $\tilde{a}_i = 0$, otherwise measures it in the Hadamard basis. Then, he computes the string $\tilde{s} = \tilde{s}_1 \tilde{s}_2 \dots \tilde{s}_{n+m}$, where $\tilde{s}_i = 0$ if the outcome of measuring $|\phi_i\rangle$ was 0 or $+$, and $\tilde{s}_i = 1$ if it was 1 or $-$.
3. For each i, Bob commits $(\tilde{s}_i, \tilde{a}_i)$ to Alice.
4. Alice chooses randomly a set of indices $T \subset \{1, \dots, n+m\}$ of size m and sends T to Bob.
5. For each $j \in T$, Bob opens the commitments associated to $(\tilde{s}_j, \tilde{a}_j)$.
6. Alice checks if $s_j = \tilde{s}_j$ whenever $a_j = \tilde{a}_j$ for all $j \in T$. If the test fails Alice aborts the protocol, otherwise she sends $a^* = a|_{\overline{T}}$ to Bob and sets $k = s|_{\overline{T}}$.
7. Bob computes $x = a^* \oplus \tilde{a}|_{\overline{T}}$ and $\tilde{k} = \tilde{s}|_{\overline{T}}$.

(Oblivious transfer phase)

8. Bob defines the two sets $I_0 = \{i \mid x_i = 0\}$ and $I_1 = \{i \mid x_i = 1\}$. Then, he sends to Alice the ordered pair $(I_c, I_{c\oplus 1})$.
9. Alice computes (e_0, e_1), where $e_i = b_i \bigoplus_{j \in I_{c\oplus i}} k_j$, and sends it to Bob.
10. Bob outputs $\tilde{b}_c = e_c \bigoplus_{j \in I_0} \tilde{k}_j$.

Figure 2. Quantum OT protocol based on secure commitments. The $\bigoplus$ denotes the bit XOR of all the elements in the family.

Intuitively, this protocol works because the computational and the Hadamard are conjugate bases. Performing a measurement in the preparation basis of a state, given by a_i, yields a deterministic outcome, whereas measuring in the conjugate basis, given by $\bar{a}_i$, results in a completely random outcome. By preparing and measuring in random bases, as shown in steps 1 and 2, approximately half of the measurement outcomes will be equal to the prepared states, and half of them will have no correlation. As Alice sends the information of preparation bases to Bob in step 6, he gets to know which of his bits are correlated with Alice's. During steps 3 to 6, Bob commits the information of his measurement basis and outcomes to Alice, who then chooses a random subset of them to test for correlations. Passing this test (statistically) ensures that Bob measured his qubits as stated in the protocol as opposed to performing a different (potentially joint) measurement. Such strategy may extract additional information from Alice's strings, but would fail to pass the specific correlation check in step 6. At step 8, Bob separates his non-tested measurement outcomes in two groups: I_0 where he measured in the same basis as the preparation one, and I_1, in which he measured in the different basis. He then inputs his bit choice c by selecting the order in which he sends the two sets to Alice. During step 9, Alice encrypts her first and second input bits with the preparation bits associated with the first and second second sets sent by Bob respectively. This effectively hides Bob's input bit because she is ignorant about the measurements that were not opened by Bob (by the security of the commitment scheme). Finally, Bob can decrypt only the bit encrypted with the preparation bits associated to I_0.

In real implementations of the protocol one should consider imperfect sources, noisy channels, and measurement errors. Thus, in step 6 Alice should perform parameter estimation for the statistics of the measurements, and pass whenever the error parameter es below some previously fixed value. Following this, Alice and Bob perform standard post-processing techniques of information reconciliation and privacy amplification before continuing to step 7. These techniques indeed work even in the presence of a dishonest Bob. As long as he has some minimal amount of uncertainty about Alice's preparation string s, an adequate privacy amplification scheme can be used to maximize Bob's uncertainty of one of Alice's input bits. This comes at the cost of increasing the amount of qubits shared per OT [32].

An example of these techniques applied in the context of the noisy storage model (where the commitment based check is replaced by a time delay under noisy memories) can be found in [19].

2.2. Oblivious Key Distribution

In order to make the quantum implementation of OT more practical during SMC we introduce the concept of oblivious keys. The protocol π_{QOT} can be separated in two phases: the *Oblivious Key Distribution phase* which consists of steps 1 to 7 and forms the π_{OKD} subprotocol, and the Oblivious Transfer phase which takes steps 8 to 10 and we denote as the $\pi_{OK\to OT}$ subprotocol. Note that after step 7 of π_{QOT} the subsets I_0, I_1 have not been revealed to Alice, so she has no information yet on how the correlated and uncorrelated bits between k and $\tilde{k}$ are distributed (recall that k and $\tilde{k}$ are the result of removing the tested bits from the strings s and $\tilde{s}$ respectively). On the other hand, after receiving Alice's preparation bases, Bob does know the distribution of correlated and uncorrelated bits between k and $\tilde{k}$, which is recorded in the string x ($x_i = 0$ if $a_i = \tilde{a}_i$, otherwise $x_i = 1$). Note that until step 7 of the protocol all computation is independent of the input bits e_0, e_1, c. Furthermore, from step 8, only the strings $k, \tilde{k}$, and x are needed to finish the protocol (in addition to the input bits). We call these three strings collectively an oblivious key, depicted in Figure 3. Formally, let Alice and Bob be two agents. Oblivious Key Distribution (OKD) is a service that outputs to Alice the string $k = k_1 k_2 \dots k_\ell$ and to Bob the string $\tilde{k} = \tilde{k}_1 \tilde{k}_2 \dots \tilde{k}_\ell$ together with the bit string $x = x_1 x_2 \dots x_\ell$, such that $k_i = \tilde{k}_i$ whenever $x_i = 0$ and $\tilde{k}_i$ does not give any information about k whenever $x_i = 1$. All of the strings are chosen at random for every invocation of the service. A pair $(k, (\tilde{k}, x))$ distributed as above is what we call an oblivious key pair. Alice, who knows k, is referred to as the sender, and Bob, who holds $\tilde{k}$ and x, is the receiver. In other words, when two parties share an oblivious key, the sender holds a string k, while the receiver has only approximately half of the bits of k, but knows exactly which of those bits he has.

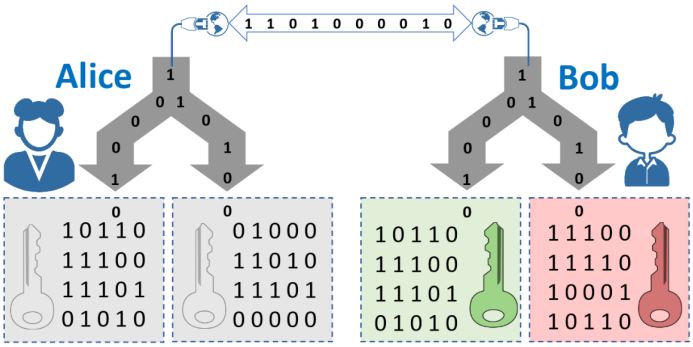


Figure 3. Oblivious keys. Alice has the string k and Bob the string $\tilde{k}$. For each party, the boxes in the left and right represent the bits of their string associated to the indices i for which x_i equals 0 (**left box**) or 1 (**right box**). Alice knows the entire key, Bob only knows half of the key, but Alice does not know which half Bob knows.

When two parties have previously shared an oblivious key pair, they can securely produce OT by performing the steps $\pi_{OK\to OT}$ of π_{QOT}. This is significantly faster than current implementations of OT without any previous shared resource and does not require quantum communication during SMC. Note that the agents can perform, previously or concurrently, an OKD protocol to share a sufficiently large oblivious key, which can be then partitioned and used to perform as many instances of OT as needed for SMC.

Fortunately, it is possible to achieve fast oblivious key exchange if the parties have access to fast and reliable quantum communications and classical commitments. In order to use this QOT protocol, the commitment scheme must be instantiated. Consider the commitment protocol π_{COMH} shown in Figure 4, first introduced by Halevi and Micali. It uses a combination of universal and cryptographic hashing, the former to ensure statistical uniformity on the commitments, and the latter to hide the

committed message. The motivation for the choice of this protocol for this task will become more apparent during the following sections as we discuss the security and efficiency characteristics of the composition of π_{QOT} with π_{COMH}, henceforth referred as the π_{HOK} (for Hybrid Oblivious Key) protocol for OT.

The existence of a reduction from OT to commitments, while proven within quantum cryptography through the π_{QOT} protocol, is an open problem in classical cryptography. The existence of commitment schemes such as π_{COMH}, which do not rely on asymmetric cryptography, provides a way to obtain OT in the quantum setting while circumventing the disadvantages of asymmetric cryptography.

Protocol π_{COMH}

Parameters: Message length $\tilde{n}$ and security parameter k. A universal hash family $\mathbf{F} = \{f : \{0,1\}^{\ell} \to \{0,1\}^{k}\}$, with $\ell = 4k + 2\tilde{n} + 4$. A collision resistant hash function H.
Parties: The verifier Alice and the committer Bob.
Inputs: Bob gets a string $\tilde{m}$ of length $\tilde{n}$.
(Commit phase)

1. Bob samples $r \in \{0,1\}^{\ell}$, computes $y = H(r)$, and chooses $f \in \mathbf{F}$, such that $f(r) = \tilde{m}$. Then, he sends (f, y) to Alice.

(Open phase)

2. Bob sends r to Alice.
3. Alice checks that $H(r) = y$. If this test fails she aborts the protocol. Otherwise, she outputs $f(r)$.

Figure 4. Commitment protocol based on collision resistant hash functions.

3. Results and Discussion

3.1. Security

In this section, we analyse the security of the proposed composition of protocols. The main result is encapsulated in the following theorem.

Theorem 1. *The protocol π_{HOK} is secure as long as the hash function is collision resistant. Moreover, if the hash function models a Random Oracle, a simple modification of the protocol can make it universally-composable secure.*

Proof. The security proof relies on several well-established results in cryptography. First, notice that the π_{HOK} protocol is closely related to the standard Quantum OT protocol π_{QOT}, which is proven statistically secure in Yao's original paper [33] and later universally composable in the quantum composability framework [34]. The difference between the two is that π_{QOT} uses ideal commitments, as opposed to the hash-based commitments in π_{HOK}. We start by showing that the protocol π_{HOK} is standalone secure. For this case, we only need to replace the ideal commitment of π_{QOT} with a standalone secure commitment protocol, such as the Halevi and Micali [29], which is depicted in π_{COMH}. Since the latter is secure whenever the hash function is collision resistant, we conclude that π_{HOK} is secure whenever the hash function is collision resistant.

Finally, we provide the simple modification of π_{HOK} that makes it universally-composable secure when the hash function models a Random Oracle. The modification is only required to improve upon the commitment protocol, as Yao's protocol with ideal commitments is universally-composable [34]. Indeed, we need to consider universally composable commitment scheme instead of π_{COMH}. This is achieved by the HMQ construction [35] which, given a standalone secure commitment scheme and a Random Oracle, outputs a universally-composable commitment scheme, which is perfectly hiding and computationally binding, that is, secure as far as collisions cannot be found. So we just need to replace π_{COMH} with the output of the HMQ construction, when π_{COMH} and H are given as inputs and H models a Random Oracle. □

Regarding the above theorem we note that, for the composable security, the HMQ construction mentioned in the proof formally requires access to a random oracle, which is an abstract object used for studying security and cannot be realized in the real world. Hence, we leave it as an additional security property, as hash functions are traditionally modelled as random oracles. Stand alone security of the π_{HOK} protocol does not require the hash function to be a random oracle.

The use of collision resistant hash functions is acceptable in the quantum setting, as it has been shown that there exist functions for which a quantum computer does not have any significant advantage in finding collisions when compared with a classical one [36]. One point to note about the security of π_{OKD} is that it is not susceptible to intercept now-decrypt later style of attacks. Bob can attempt an attack in which he does not properly measure the qubits sent by Alice at step 2, and instead waits until Alice reveals the test subset in step 4 to measure honestly only those qubits. For that he must be able to control the openings of the commitment scheme such that Alice opens the values of his measurement outcomes for those qubits. In order to do this, he must be able find collisions for H before step 5. This means that attacking the protocol by finding collisions of the hash function is only effective if it is done in real time, that is, between steps 3 and 5 of the protocol. This is in contrast to asymmetric cryptography based OT, in which Bob can obtain both bits if he is able to overcome the computational security at a later stage.

Finally, we point out that the OT extension algorithms that are used during SMC often rely only on collision resistant hash functions [37] anyway. If those protocols are used to extend the base OTs produced by π_{HOK}, we can effectively speed up the OT rates without introducing any additional computational complexity assumption.

3.2. *Efficiency*

Complexity-wise, the main problem with public-key based OT protocols is that they require a public/private key generation, encryption, and decryption per transfer. In the case of RSA and ElGamal based algorithms, this has complexity $O(n^{2.58})$ (where $N = 2^n$ is the size of the group), using Karatsuba multiplication and Berett reduction for Euclidian division [38]. Post-quantum protocols are still ongoing optimization, but recent results show RLWE key genereration and encryption in time $O(n^2 \log(n))$ [39].

To study the time complexity of the π_{HOK} protocol, consider first the complexity of π_{COMH}. It requires two calls of H and one call of the universal hash family $\mathbf{F}$, $\tilde{n}$ bit comparisons (if using the technique proposed in [29] to find the required f), and one additional evaluation of f. Cryptographic hash functions are designed so that their time complexity is linear on the size of the input, which in this case is $\ell = 4k + 2\tilde{n} + 4$. To compute the universal hashing, the construction in [29] requires $\tilde{n}k$ binary multiplications. Thus, the running time of π_{COMH} is linear on the security parameter k. On the other hand, π_{QOT} has two security parameters: n, associated to the size of the keys used to encrypt the transferred bits, and m, associated to the security of the measurement test done by Alice. The protocol requires $n + m$ qubit preparations and measurements, $n + m$ calls of the commitment scheme, and n bit comparisons. This leads to an overall time complexity of $O(k(n + m))$ for the π_{HOK} protocol, which is linear in all of its security parameters.

In realistic scenarios, however, error correction and privacy amplification must be implemented during the $\pi_{OK \to OT}$. For the former, LDPC codes [40] or the cascade algorithm [41] can be used, and the latter can be done with universal hashing. For a given channel error parameter, these algorithms have time complexity linear in the size of the input string, which in our case is n. Hence, π_{HOK} stays efficient when considering channel losses and preparation/measurement errors.

One of the major bottlenecks in the GMW protocol for SMC is the number of instances of OT required (it is worth noting that GMW uses 1-out-of-4 OT, which can efficiently be obtained from two instances of the 1-out-of-2 OT presented here [42]). A single Advanced Encryption Standard (AES) circuit can be obtained with the order of 10^6 instances of OT. However, with current solutions, i.e., with computational implementations of OT based on asymmetric classical cryptography, one can

generate $\sim 10^3$ secure OTs per second in standard devices [43]. It is possible to use OT extension algorithms to increase its size up to rates of the order of 10^6 OT per second [3]. Several of such techniques are based on symmetric cryptography primitives [43], such as hash functions, and could also be used to extend the OTs generated by π_{HOK}.

Due to the popularity of crypto-currencies, fast and efficient hashing machines have recently become more accessible. Dedicated hashing devices are able to compute SHA-256 at rates of 10^{12} hashes per second (see Bitfury, Ebit, and WhatsMiner, for example). In addition, existent standard Quantum Key Distribution (QKD) setups can be adapted to implement OKD, since both protocols share the same requirements for the generation and measurement of photons. Notably, QKD setups have already demonstrated secret key rates of the order of 10^6 bits per second [44–48]. It is also worth mentioning that, as opposed to QKD, OKD is useful even in the case when Alice and Bob are at the same location. This is because in standard key distribution the parties trust each other and, if at the same location, they can just exchange hard drives with the shared key, whereas when sharing oblivious keys, the parties do not trust each other and need a protocol that enforces security. Thus, for the cases in which both parties being at the same location is not an inconvenience, the oblivious key rates can be further raised, as the effects of channel noise are minimized.

Direct comparisons of OT generation speed between asymmetric cryptography techniques and quantum techniques are difficult because the algorithms run on different hardware. Nevertheless, as quantum technologies keep improving, the size and cost of devices capable of implementing quantum protocols will decrease and their use can result in significant improvements of OT efficiency, in the short-to-medium term future.

4. Conclusions

Motivated by the usefulness of SMC as a privacy-protecting data mining tool, and identifying its OT cost as its main implementation challenge, we have proposed a potential solution for practical implementation of OT as a subroutine SMC. The scheme consists on pre-sharing an oblivious key pair and then using it to compute fast OT during the execution of the SMC protocol. We call this approach hybrid because it uses resources traditionally associated with classical symmetric cryptography (cryptographic hash functions), as well as quantum state communication and measurements on conjugate observables, resources associated with quantum cryptography. The scheme is secure as far as the chosen hash function is secure against quantum attacks. In addition, we showed that the overall time complexity of π_{HOK} is linear on all its security parameters, as opposed to the public-key based alternatives, whose time complexities are at least quadratic on their respective parameters. Finally, by comparing the state of current technology with the protocol requirements, we concluded that it has the potential to surpass current asymmetric cryptography based techniques.

It was also noted that current experimental implementations of standard discrete-variable QKD can be adapted to perform π_{HOK}. The same post-processing techniques of error correction and privacy amplification apply, however, fast hashing subroutines should be added for commitments during the parameter estimation step. Future work includes designing an experimental setup, meeting the implementation challenges, and experimentally testing the speed, correctness, and security of the resulting oblivious key pairs. This includes computing oblivious key rate bounds for realistic scenarios and comparing them with current alternative technologies. Real world key rate comparisons can help us understand better the position of quantum technologies in the modern cryptographic landscape.

Regarding the use of quantum cryptography during the commitment phase; because of the impossibility theorem for unconditionally secure commitments in the quantum setting [17], one must always work with an additional assumption on top of needing quantum resources. The noisy storage model provides an example in which the commitments are achieved by noisy quantum memories [21,22,49]. The drawback of this particular assumption is the fact that advances in quantum storage technology work against the performance of the protocol, which is not a desired feature. The added cost of using quantum communication is a disadvantage. So far, to the knowledge of the

authors, there are no additional practical quantum bit commitment protocols that provide advantages in security or efficiency compared to classical ones once additional assumptions (such as random oracles, common reference strings, computational hardness, etc.,) are introduced. Nevertheless, we are optimistic that such protocols can be found in the future, perhaps by clever design, or by considering a different a kind of assumption outside of the standard ones.

Author Contributions: Conceptualization, P.M. and A.N.P.; methodology, M.F.R., N.A.S. and N.J.M.; validation, M.L., N.P., and A.S.; formal analysis, M.L., P.Y., N.P., A.S. and P.M.; investigation, M.L., M.F.R., N.A.S. and N.J.M.; writing—original draft preparation, M.L., N.P., P.Y., M.F.R., N.A.S., N.J.M. and A.N.P.; writing—review and editing, M.L. and P.M.; visualization, M.F.R., N.A.S., N.J.M. and A.N.P.; supervision, P.M. and A.N.P.; project administration, P.M., A.S. and A.N.P.; funding acquisition, P.M., A.S. and A.N.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Fundação para a Ciência e a Tecnologia (FCT) through national funds, by FEDER, COMPETE 2020, and by Regional Operational Program of Lisbon, under UIDB/50008/2020, UIDP/50008/2020, UID/CEC/00408/2013, POCI-01-0145-FEDER-031826, POCI-01-0247-FEDER-039728, PTDC/CCI-CIF/29877/2017, PD/BD/114334/2016, PD/BD/113648/2015, and CEECIND/04594/2017/CP1393/CT0006A.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Lindell, Y.; Pinkas, B. Secure Multiparty Computation for Privacy-Preserving Data Mining. *J. Priv. Confid.* **2009**, 59–98. [CrossRef]
2. Laud, P.; Kamm, L. *Applications of Secure Multiparty Computation*; IOS Press: Amsterdam, The Netherlands, 2015; Volume 13.
3. Asharov, G.; Lindell, Y.; Schneider, T.; Zohner, M. More Efficient Oblivious Transfer Extensions. *J. Cryptol.* **2017**, *30*, 805–858. [CrossRef]
4. Bernstein, D.J.; Lange, T. Post-quantum cryptography. *Nature* **2017**, *549*, 188. [CrossRef] [PubMed]
5. Yao, A.C.C. How to generate and exchange secrets. In Proceedings of the 27th Annual Symposium on Foundations of Computer Science (sfcs 1986), Toronto, ON, Canada, 27–29 October 1986; pp. 162–167. [CrossRef]
6. Goldreich, O.; Micali, S.; Wigderson, A. How to Play ANY Mental Game. In Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing, New York, NY, USA, 25–27 May 1987; ACM: New York, NY, USA, 1987; pp. 218–229. [CrossRef]
7. Schneider, T.; Zohner, M., GMW vs. Yao? Efficient Secure Two-Party Computation with Low Depth Circuits. In *Financial Cryptography and Data Security, Proceedings of the 17th International Conference, FC 2013, Okinawa, Japan, 1–5 April 2013*; Revised Selected Papers; Sadeghi, A.R., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 275–292.
8. Beaver, D.; Micali, S.; Rogaway, P. The round complexity of secure protocols. In Proceedings of the Twenty-Second Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, 14–16 May 1990; pp. 503–513.
9. Kilian, J. Founding Cryptography on Oblivious Transfer. In Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, 2–4 May 1988; ACM: New York, NY, USA, 1988; pp. 20–31.
10. Harnik, D.; Ishai, Y.; Kushilevitz, E. How many oblivious transfers are needed for secure multiparty computation? In Proceedings of the Annual International Cryptology Conference, Santa Barbara, CA, USA, 19–23 August 2007; pp. 284–302.
11. Rabin, M.O. *How To Exchange Secrets*; Technical Report TR-81; Aiken Computation Laboratory, Harvad University, Cambridge, MA, USA, 1981.
12. Goldreich, O. *Foundations of Cryptography, Volume I Basic Techniques*; Cambridge University Press: Cambridge, UK, 2001.
13. Paar, C.; Pelzl, J. *Understanding Cryptography*; Springer: Berlin/Heidelberg, Germany, 2010.

14. Broadbent, A.; Schaffner, C. Quantum cryptography beyond quantum key distribution. *Des. Codes Cryptogr.* **2016**, *78*, 351–382. [CrossRef]
15. Shenoy-Hejamadi, A.; Pathak, A.; Radhakrishna, S. Quantum cryptography: Key distribution and beyond. *Quanta* **2017**, *6*, 1–47. [CrossRef]
16. Lo, H.K.; Chau, H.F. Is Quantum Bit Commitment Really Possible? *Phys. Rev. Lett.* **1997**, *78*, 3410–3413. [CrossRef]
17. Mayers, D. Unconditionally Secure Quantum Bit Commitment is Impossible. *Phys. Rev. Lett.* **1997**, *78*, 3414–3417. [CrossRef]
18. Wehner, S.; Schaffner, C.; Terhal, B.M. Cryptography from Noisy Storage. *Phys. Rev. Lett.* **2008**, *100*, 220502. [CrossRef]
19. Wehner, S.; Curty, M.; Schaffner, C.; Lo, H.K. Implementation of two-party protocols in the noisy-storage model. *Phys. Rev. A* **2010**, *81*, 052336. [CrossRef]
20. Konig, R.; Wehner, S.; Wullschleger, J. Unconditional Security From Noisy Quantum Storage. *IEEE Trans. Inf. Theory* **2012**, *58*, 1962–1984. [CrossRef]
21. Loura, R.; Almeida, Á.J.; André, P.; Pinto, A.; Mateus, P.; Paunković, N. Noise and measurement errors in a practical two-state quantum bit commitment protocol. *Phys. Rev. A* **2014**, *89*, 052336. [CrossRef]
22. Almeida, Á.J.; Stojanovic, A.D.; Paunković, N.; Loura, R.; Muga, N.J.; Silva, N.A.; Mateus, P.; André, P.S.; Pinto, A.N. Implementation of a two-state quantum bit commitment protocol in optical fibers. *J. Opt.* **2015**, *18*, 015202. [CrossRef]
23. Erven, C.; Ng, N.; Gigov, N.; Laflamme, R.; Wehner, S.; Weihs, G. An experimental implementation of oblivious transfer in the noisy storage model. *Nat. Commun.* **2014**, *5*, 3418. [CrossRef] [PubMed]
24. Furrer, F.; Gehring, T.; Schaffner, C.; Pacher, C.; Schnabel, R.; Wehner, S. Continuous-variable protocol for oblivious transfer in the noisy-storage model. *Nat. Commun.* **2018**, *9*, 1450. [CrossRef] [PubMed]
25. Ng, N.H.Y.; Joshi, S.K.; Chen Ming, C.; Kurtsiefer, C.; Wehner, S. Experimental implementation of bit commitment in the noisy-storage model. *Nat. Commun.* **2012**, *3*, 1326. [CrossRef] [PubMed]
26. Lunghi, T.; Kaniewski, J.; Bussières, F.; Houlmann, R.; Tomamichel, M.; Wehner, S.; Zbinden, H. Practical Relativistic Bit Commitment. *Phys. Rev. Lett.* **2015**, *115*, 030502. [CrossRef] [PubMed]
27. Verbanis, E.; Martin, A.; Houlmann, R.; Boso, G.; Bussières, F.; Zbinden, H. 24-Hour Relativistic Bit Commitment. *Phys. Rev. Lett.* **2016**, *117*, 140506. [CrossRef]
28. Pitalúa-García, D.; Kerenidis, I. Practical and unconditionally secure spacetime-constrained oblivious transfer. *Phys. Rev. A* **2018**, *98*, 032327. [CrossRef]
29. Halevi, S.; Micali, S. Practical and Provably-Secure Commitment Schemes from Collision-Free Hashing. In *Advances in Cryptology—CRYPTO '96, Proceedings of the 16th Annual International Cryptology Conference Santa Barbara, California, CA, USA, 18–22 August 1996*; Koblitz, N., Ed.; Springer: Berlin/Heidelberg, Germany, 1996; pp. 201–215.
30. Pinto, A.N.; Ramos, M.F.; Silva, N.A.; Muga, N.J. Generation and Distribution of Oblivious Keys through Quantum Communications. In Proceedings of the 2018 20th International Conference on Transparent Optical Networks (ICTON), Bucharest, Romania, 1–5 July 2018; pp. 1–3. [CrossRef]
31. Ramos, M.F.; Silva, N.A.; Muga, N.J.; Pinto, A.N. Reversal operator to compensate polarization random drifts in quantum communications. *Opt. Express* **2020**, *28*, 5035–5049. [CrossRef]
32. Lindell, Y.; Pinkas, B. An efficient protocol for secure two-party computation in the presence of malicious adversaries. In Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques, Barcelona, Spain, 20–24 May 2007; pp. 52–78.
33. Yao, A.C.C. Security of Quantum Protocols Against Coherent Measurements. In Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing, Las Vegas, NV, USA, 29 May–1 June 1995; ACM: New York, NY, USA, 1995; pp. 67–75.
34. Unruh, D. Universally composable quantum multi-party computation. In Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques, French Riviera, France, 30 May–3 June 2010; pp. 486–505.
35. Hofheinz, D.; Müller-Quade, J. Universally Composable Commitments Using Random Oracles. In *Theory of Cryptography*; Naor, M., Ed.; Springer: Berlin/Heidelberg, Germany, 2004; pp. 58–76.
36. Aaronson, S.; Shi, Y. Quantum Lower Bounds for the Collision and the Element Distinctness Problems. *J. ACM* **2004**, *51*, 595–605. [CrossRef]

37. Asharov, G.; Lindell, Y.; Schneider, T.; Zohner, M. More Efficient Oblivious Transfer and Extensions for Faster Secure Computation. In Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security, Berlin, Germany, 4–8 November 2013; ACM: New York, NY, USA, 2013; pp. 535–548. [CrossRef]
38. Menezes, A.J.; Katz, J.; Van Oorschot, P.C.; Vanstone, S.A. *Handbook of Applied Cryptography*; CRC Press, Inc.: Boca Raton, FL, USA: 1996.
39. Ding, J.; Xie, X.; Lin, X. A Simple Provably Secure Key Exchange Scheme Based on the Learning with Errors Problem. *Iacr Cryptol. Eprint Arch.* **2012**, *2012*, 688.
40. Martinez-Mateo, J.; Elkouss, D.; Martin, V. Key reconciliation for high performance quantum key distribution. *Sci. Rep.* **2013**, *3*, 1576. [CrossRef] [PubMed]
41. Brassard, G.; Salvail, L. Secret-key reconciliation by public discussion. In *Workshop on the Theory and Application of of Cryptographic Techniques*; Springer: Berlin/Heidelberg, Germany, 1993; pp. 410–423.
42. Naor, M.; Pinkas, B. Computationally secure oblivious transfer. *J. Cryptol.* **2005**, *18*, 1–35. [CrossRef]
43. Chou, T.; Orlandi, C. The simplest protocol for oblivious transfer. In Proceedings of the International Conference on Cryptology and Information Security in Latin America, Guadalajara, Mexico, 23–26 August 2015; pp. 40–58.
44. Comandar, L.; Fröhlich, B.; Lucamarini, M.; Patel, K.; Sharpe, A.; Dynes, J.; Yuan, Z.; Penty, R.; Shields, A. Room temperature single-photon detectors for high bit rate quantum key distribution. *Appl. Phys. Lett.* **2014**, *104*, 021101. [CrossRef]
45. Islam, N.T.; Lim, C.C.W.; Cahall, C.; Kim, J.; Gauthier, D.J. Provably secure and high-rate quantum key distribution with time-bin qudits. *Sci. Adv.* **2017**, *3*, e1701491. [CrossRef]
46. Ko, H.; Choi, B.S.; Choe, J.S.; Kim, K.J.; Kim, J.H.; Youn, C.J. High-speed and high-performance polarization-based quantum key distribution system without side channel effects caused by multiple lasers. *Photonics Res.* **2018**, *6*, 214–219. [CrossRef]
47. Wang, T.; Huang, P.; Zhou, Y.; Liu, W.; Ma, H.; Wang, S.; Zeng, G. High key rate continuous-variable quantum key distribution with a real local oscillator. *Opt. Express* **2018**, *26*, 2794–2806. [CrossRef]
48. Pirandola, S.; Andersen, U.; Banchi, L.; Berta, M.; Bunandar, D.; Colbeck, R.; Englund, D.; Gehring, T.; Lupo, C.; Ottaviani, C.; et al. Advances in Quantum Cryptography. *arXiv* **2019**, arXiv:1906.01645.
49. Loura, R.; Arsenović, D.; Paunković, N.; Popović, D.B.; Prvanović, S. Security of two-state and four-state practical quantum bit-commitment protocols. *Phys. Rev. A* **2016**, *94*, 062335. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

entropy

MDPI

Article

Information-Theoretical Criteria for Characterizing the Earliness of Time-Series Data

Mariano Lemus [1,2], João P. Beirão [1,2], Nikola Paunković [1,2], Alexandra M. Carvalho [1,2,*] and Paulo Mateus [1,2]

1 Instituto de Telecomunicações, 1049-001 Lisboa, Portugal
2 Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal
* Correspondence: alexandra.carvalho@lx.it.pt

Received: 2 December 2019; Accepted: 26 December 2019; Published: 30 December 2019



Abstract: Biomedical signals constitute time-series that sustain machine learning techniques to achieve classification. These signals are complex with measurements of several features over, eventually, an extended period. Characterizing whether the data can anticipate prediction is an essential task in time-series mining. The ability to obtain information in advance by having early knowledge about a specific event may be of great utility in many areas. Early classification arises as an extension of the time-series classification problem, given the need to obtain a reliable prediction as soon as possible. In this work, we propose an information-theoretic method, named Multivariate Correlations for Early Classification (MCEC), to characterize the early classification opportunity of a time-series. Experimental validation is performed on synthetic and benchmark data, confirming the ability of the MCEC algorithm to perform a trade-off between accuracy and earliness in a wide-spectrum of time-series data, such as those collected from sensors, images, spectrographs, and electrocardiograms.

Keywords: Akaike information criterion; minimum description length; time-series charaterization

1. Introduction

A time-series (TS) consists of measurements or observations acquired and organized sequentially over time. In this context, one or multiple variables may be examined, being the first-named univariate time-series (UTS) and the second multivariate time-series (MTS). Several data mining application areas deliver this sort of data, such as medicine, economy, meteorology, and marketing. Standard TS classification involves using temporal data for constructing a classifier, which can predict the class label of a new given TS, with satisfactory accuracy.

Early classification (EC) is an extension of the TS classification problem, and it arises in scenarios where the anticipation of the prediction is beneficial. This matter has recently been a relevant subject of study, due to its several time-sensitive applications. For instance, a medical study [1] described how clinical data revealed that infants diagnosed with sepsis disease suffered from an unusual heartbeat twenty-four hours before the diagnosis. In this case, supervising the TS data of the infant's heartbeat and being able to classify it in advance may lead to effective early diagnosis and treatment.

The work from Xing et al. [2] was one of the first to formulate the problem of EC, proposing to unveil a timestamp from which the information of the TS from that point on is irrelevant. As stated by Xing et al. [3], it is vital to distinguish EC from classic TS prediction, where the goal is to forecast values given the whole TS. EC of temporal data consists of anticipating the classification by using only a portion of the available information, without compromising the prediction quality. There are two essential requirements for an early classifier: being able to designate the earliest time location

(timestamp) of accurate classification, and ensuring an accuracy close to the case of using the full-length TS.

Several methods addressing the early classification problem have been proposed in the last years [2,4–8]. These approaches require a learning stage followed by a classification step, being able to assign a class label to a single incomplete time-series. Notwithstanding the significant advances in EC, there are temporal datasets for which it is intrinsically hard to perform early predictions, due to phenomenons like high conductance of the underlying Markov model, or cumulative errors throughout the time measurements.

In contrast, we propose a dataset investigation, where the information from the entire collection of time-series (and not only a single TS), as well as their respective class labels, is the subject of study. Our approach explores the knowledge contained in the data via the temporal correlation of the variables. In general, testing all possible timestamps without external information would require an expensive, and in some cases, computationally unfeasible, cross-validation search. We propose to use information-theoretical criteria to perform this analysis efficiently throughout the whole data.

Overall, the main contribution of this work is applying information-theoretic criteria for examining the EC opportunity in a dataset containing univariate or multivariate TS, extracting a global timestamp (for the overall data) from which accurate prediction might be performed. An EC method, like those mentioned before, can then be applied, with a priori knowledge of the EC opportunity in the data.

We assessed the merits of the proposed method in a wide-spectrum of time-series data, such as those collected from sensors, images, spectrographs, and electrocardiograms. In particular, we found that for a collection of time-series that traces the electrical activity recorded during one heartbeat, the first 30% of the signal contains enough information to discriminate between a normal heartbeat and a myocardial infarction, leveraging the development of novel methods to focus in this fraction of the signal. An implementation of the proposed criteria, through a procedure called Multivariate Correlations for Early Classification (MCEC), is made freely available to the comunity at https://joaopbeirao.github.io/MCECalgorithm/.

The paper is organized as follows. In Section 2, we review some basic concepts of Bayesian networks and multivariate correlation. In Section 3, we present our EC opportunity algorithm, followed by experimental results over benchmark data; results in synthetic data are left for the Appendices. Finally, we draw some conclusions and discuss future work.

2. Background

We start by introducing Bayesian networks as they have a well-established framework to measure model complexity. In this regard, we choose two criteria: minimum description length and Akaike information criterion. Finally, we present multivariate correlations to assess data earliness by deriving the dependencies among the variables through a concrete Bayesian network calibrated with the adopted criteria.

2.1. Bayesian Networks

Probabilistic graphical models attempt to describe the behaviour of complex systems using a graph-based framework for representing the probability distributions [9]. Bayesian networks (BNs) are probabilistic graphical models for describing complex domains, and they can be used to represent the information about an uncertain system [10–12]. The BN representation consists of a directed acyclic graph G, characterized by a set of nodes $\mathcal{N} = \{X_1, X_2, \ldots, X_n\}$ and a set of directed edges E. Considering a $G = (\mathcal{N}, E)$, each node (vertex) corresponds to a random variable X_i, and the edges (arrows), that connect the nodes in a specific direction, describe the probabilistic dependencies between the random variables. For each node X_i, two sets can be defined: the set of parents Π_{X_i} and the set of non-descendants Φ_{X_i}. The structure of a BN is based on the assumption that each node X_i is conditionally independent of Φ_{X_i}, provided that the values of the variable in Π_{X_i} are known. The group of local probability models, representing the dependence of each variable X_i on Π_{X_i}, specifies

the parameters for describing the network structure. These form the set of conditional probability distributions

$$\Theta = \left\{\theta_{X_i|\Pi_{X_i}}\right\}_{i\in\{1,\ldots,n\}},$$

where $\theta_{X_i|\Pi_{X_i}} = P(X_i = x_i \mid \Pi_{X_i} = \omega_i)$, associated to each node X_i and conditioned on Π_{X_i}.

A BN $\mathcal{B} = (G, \Theta)$ is comprised of the direct acyclic graph structure G together with the set of parameters Θ. The joint probability distribution defined by this representation is calculated as:

$$P_{\mathcal{B}}(X_1, \ldots, X_n) = \prod_{i=1}^{n} P_{\mathcal{B}}(X_i \mid \Pi_{X_i}) = \prod_{i=1}^{n} \theta_{X_i|\Pi_{X_i}}. \tag{1}$$

For a given a multinomial dataset D of size N, the problem of learning a BN consists of designing the $\mathcal{B} = (G, \Theta)$ that best represents D, according to some scoring function [13]. This scoring criterion corresponds to the search guide for evaluating the effectiveness of the network in representing the data. Moreover, when the structure of the network is fixed, the parameters Θ that optimize the likelihood, for a given dataset, are those described by the observed frequency estimates:

$$\hat{P}_{\mathcal{B}}(X_i = x_i \mid \Pi_{X_i} = \omega_i) = \frac{|D_{x_i,\omega_i}|}{|D_{\omega_i}|}, \tag{2}$$

for which $|D_{x_i,\omega_i}|$ represents the number of instances in D, where X_i takes the value x_i, and its parents Π_{X_i} take the value ω_i. Similarly, $|D_{\omega_i}|$ denotes the number of instances in D, where Π_{X_i} takes the value ω_i.

The Minimum Description Length (MDL) principle is known as an Occam's razor approach to select, for a given dataset, the best fitting model and its parameters [14]. It is a widely used metric which states that, for a certain data and a number of alternative models, the best option corresponds to the simplest model [15–17]. In the problem of learning a BN, the Bayesian Information Criterion (BIC) is known as the MDL score. It is concerned with analysing the trade-off between the Log-Likelihood (LL) of the dataset D (the effectiveness of the fit to the data) and the complexity of the model. This scoring function [13] is defined as

$$MDL(D \mid \mathcal{B}) = -LL(D \mid \mathcal{B}) + \frac{\log_2 N}{2}|\mathcal{B}|, \tag{3}$$

where N corresponds to the size of the data, and $|\mathcal{B}|$ represents the model dimension (number of parameters in $\mathcal{B}$). The LL term quantifies the amount of information required to describe the dataset D, using $\mathcal{B}$. Conversely, the penalty term measures the amount of information needed to encode the model $\mathcal{B}$. The overall goal of the MDL score is to elicit the model that most effectively fits the dataset, provided that its complexity is as low as possible, in this way avoiding overfitting.

Similarly to the MDL scoring function, the Akaike Information Criterion (AIC) [18] corresponds to a measure of the quality of statistical models for describing a given dataset. In the problem of learning a BN, the difference between MDL and AIC is associated to the penalty applied to the number of parameters $|\mathcal{B}|$. The AIC scoring function [13] can be defined as:

$$AIC(D \mid \mathcal{B}) = -LL(D \mid \mathcal{B}) + |\mathcal{B}|. \tag{4}$$

In Equation (3), the second term quantifies the amount of information required to encode the model $\mathcal{B}$, where each parameter in the set Θ is considered to use $\frac{1}{2}\log_2 N$ bits. Conversely, in Equation (4) each parameter of Θ is considered to use 1 bit. This means that the penalization on the number of independent parameters is stronger in the MDL scoring function than in the AIC score. Likewise for the MDL score, the best model corresponds to the one that minimizes Equation (4).

Literature complies with the fact that these two criteria demonstrate different properties for model selection and that they are appropriate according to specific conditions [19–21]. According to Vrieze [21], MDL is considered to be consistent in selecting the true model, with probability close to one, given that the true model is in the set of candidate models. On the other hand, if the true model is not in the set of alternatives, AIC is considered to be effective, since it selects the model that minimizes the mean squared error of the estimation. However, both criteria are unsuitable for dealing with low dimensional datasets for which the number of instances is close to the number of parameters to estimate [22].

2.2. Multivariate Correlations

We briefly introduce multivariate correlations to derive a concrete Bayesian network from which we deduce model complexity. From a statistical point of view, the concept of correlation between variables attempts to measure the relationships and dependencies among them. The knowledge of how the variables are related, as well as of what inferences can be made about their causal relationships, is useful for drawing conclusions about potential predictive relationships to be analyzed and exploited.

For a finite set of discrete random variables $S = \{X_i\}_{i=1,\dots,n}$, with a joint probability distribution $P_S(X_1, \dots, X_n)$, the total correlations between those variables can be defined as [23]:

$$I(X_1, \dots, X_n) = \sum_{i=1}^{n} H(X_i) - H(X_1, \dots, X_n). \tag{5}$$

In this case, the mutual information measures the dependencies among the variables, i.e., the amount of information that these quantities give about each other.

Let a structural relation R be a subset of the system S. Its joint probability distribution corresponds to the marginal distribution from S

$$P_R(X_{R_1}, \dots, X_{R_k}) = \sum_{X_i \notin R} P_S(X_1, \dots, X_n), \tag{6}$$

where k is the number of elements in R.

Following the maximum entropy principle applied to fixed marginals [24], we say that a structure associated to the system S with underlying joint probability P_S is a pair $(\mathcal{S}, P_{\mathcal{S}})$, where $\mathcal{S} = \{R_j\}_{j=1,\dots,k}$ is a collection of structural relations and $P_{\mathcal{S}}$ is (another) joint probability distribution over S, such that:

1. No $R_i \in \mathcal{S}$ is contained in another $R_j \in \mathcal{S}$ (or in other words, $\forall_{i \neq j}\ R_i \nsubseteq R_j$);
2. Every $X_i \in S$ is included in at least one $R_j \in \mathcal{S}$;
3. $P_{\mathcal{S}}$ is the solution to the optimization problem:

$$\max_{P \in \mathcal{P}} \ H(P) \ \text{ s.t. } \sum_{X_i \notin R_j} P_{\mathcal{S}}(X_1, \dots, X_n) = \sum_{X_i \notin R_j} P_S(X_1, \dots, X_n), \forall R_j \in \mathcal{S},$$

where $\mathcal{P}$ is the set of probability distributions of the variables from S.

For example, from the set of discrete random variables $S = \{X_1, X_2, X_3, X_4\}$, some admissible structures $\mathcal{S}$ are $\{\{X_1, X_2, X_3\}, \{X_4\}\}$, $\{\{X_1, X_2\}, \{X_3, X_4\}\}$ or $\{\{X_1, X_2\}, \{X_1, X_4\}, \{X_2, X_3, X_4\}\}$. On the other hand, the structure

$$\mathcal{S} = \{\{X_1, X_2, X_3\}, \{X_1, X_3\}, \{X_4\}\}$$

is not an acceptable structure since the relation between X_1 and X_3 is included in two structural relations, which violates the first condition. Similarly, $\mathcal{S} = \{\{X_1, X_2\}, \{X_2, X_4\}\}$ does not consist of a proper structure because the variable $X_3 \in S$ is not part of any structural relation from $\mathcal{S}$, as required by the second condition.

For a given system $S = \{X_i\}_{i=1,\dots,n}$ and an associated set of structural relations $\mathscr{S} = \{R_j\}_{j=1,\dots,k}$, the mutual information $I(S)$ represents the maximum amount of information that the variables X_i from S provide about each other. On the other hand, $I(\mathscr{S})$ quantifies the information described by the correlations only inside the structural relations R_j. The difference $I(S) - I(\mathscr{S})$ measures the knowledge of the dependencies and relationships between the variables of S that are not included in the relations that compose $\mathscr{S}$. From Equation (5), this value can be described as a difference of entropies:

$$I(S) - I(\mathscr{S}) = \sum_{i=1}^{n} H(X_i) - H(S) - \sum_{i=1}^{n} H(X_i) - H(\mathscr{S}) = H(\mathscr{S}) - H(S). \tag{7}$$

The above expression is always non-negative because the distribution of $\mathscr{S}$ has the maximum entropy in a set where the distribution of S belongs. Seeing that the entropy quantifies the average uncertainty of a random variable, $H(\mathscr{S}) - H(S)$ represents the information given by the existing correlations in S, that is not incorporated in the structural relations from $\mathscr{S}$.

3. Proposed Method

We now set out to derive the proposed EC method for temporal data. For convenience, we start by introducing a few additional notations. Let a dataset D be a collection of pairs (T_j, c_j) for all $j \in \{1, \dots, N\}$, where T_j consists of a TS, c_j corresponds to its respective class label and N is the number of instances in D. In general, a TS is defined as a vector of length L

$$T_j = \left(\mathbf{x}_1^{(j)}, \mathbf{x}_2^{(j)}, \dots, \mathbf{x}_L^{(j)}\right), \tag{8}$$

where each component $\mathbf{x}_k^{(j)} = \left(x_{k1}^{(j)}, x_{k2}^{(j)}, \dots, x_{km}^{(j)}\right)$ consists of m features measured at *time point* (TP) $k \in \{1, \dots, L\}$. The object of TS classification is to associate a class label c to a given time-series T (not necessarily in the data).

Consider a TS T, as in Equation (8), representing the evolution of the random vector $\mathbf{X} = (X_1, \dots X_m)$ over time, and its respective class label, which is denoted by the random variable C. The set of $\mathbf{X}_k$ can be viewed as a collection of time-dependent discrete random variables, for which a joint probability distribution can be defined. Note that, since a TS is chronologically organized, it is relevant to analyze the dependence of variables on their early states, i.e., the degree of dependence of $\mathbf{X}$ at a certain TP on the value observed at a previous instant. Similarly, the correlation between C and $\mathbf{X}_k$ quantifies the dependence that the vector $\mathbf{X}$ at TP k has on the class label. In the EC context, the focus is to study systems where the class labels verify a high dependence on a certain amount of early states of $\mathbf{X}_k$, while the remaining TPs are dispensable for a satisfactory classification.

Consider the finite set of discrete random variables S to be composed of the TS T together with its respective class label C. The system

$$S = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n, \mathbf{X}_{n+1}, \dots, \mathbf{X}_L, C\}$$

has an associated joint probability distribution $P_S(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_L, C)$, where L represents the TS length. The goal is to find the value n and the distribution P_S^n, such that $P_S^n(C \mid \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) \approx P_S(C \mid \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_L)$. Therefore, $P_S^n(C \mid \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$ and $P_S(C \mid \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_L)$ describe the probability of the class label C occurring, provided that the first n, or all variables of T, are known, respectively.

In some cases, the joint probability distribution P_S is not known in advance; thus, it has to be computed from the data, through maximum likelihood estimation. In particular, for a dataset D of size N, the distribution P_S that maximizes the likelihood of D is such that

$$\hat{P}_S(\mathbf{X}_1 = \mathbf{x}_1, \dots \mathbf{X}_L = \mathbf{x}_L, C = c) = \frac{|D_{\mathbf{x}_1,\dots,\mathbf{x}_L,c}|}{N}, \tag{9}$$

where $|D_{\mathbf{x}_1,\dots,\mathbf{x}_L,c}|$ is the number of instances in D for which $\mathbf{X}_i$ takes the value $\mathbf{x}_i$ and C the value c.

Given the system S, the set of structural relations, defined by

$$\mathcal{S}_n = \{\{\mathbf{X}_1, \ldots, \mathbf{X}_n, \mathbf{X}_{n+1}, \ldots, \mathbf{X}_L\}, \{\mathbf{X}_1, \ldots, \mathbf{X}_n, C\}\},$$

depends on the value of n and it corresponds to a structure that respects the previously described properties. Considering the sets $A_n = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ and $B_n = \{\mathbf{X}_{n+1}, \ldots, \mathbf{X}_L\}$, the structure is represented as $\mathcal{S}_n = \{\{A_n, B_n\}, \{A_n, C\}\}$. The structural relation A_n contains information about the evolution of the variable $\mathbf{X}$ until the TP n, i.e., the early stages of the collection of TS. On the other hand, B_n describes the remaining instants of T_i, which can be viewed as the knowledge about the later stages of the variable $\mathbf{X}$. Finally, C represents the class label information from the collection of TS. The structure $\mathcal{S}_n$ can be seen as a simplified model of the system S. It is expected to include the correlations between the early and the later information about the TS (A_n and B_n), as well as between the early states of T_i and the knowledge about their classes (A_n and C). Conversely, the correlations between B_n and C are not preserved because the idea is to study the possibility of describing the class from the early states A_n, while neglecting the information from B_n. The probability distribution of $\mathcal{S}_n$ is obtained based on Theorem 1 (the proof can be found in Appendix A) and considering the BN represented in Figure 1.

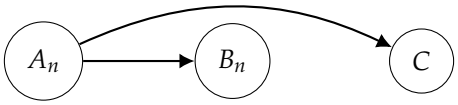


Figure 1. Bayesian network representation of the structure $\mathcal{S}_n$ from the system S.

Theorem 1. *Consider the Bayesian network $\mathcal{B}_n = (G_n, \Theta_n)$ with G_n given by Figure 1 and Θ_n calculated according to Equation (2). The structure $(\mathcal{S}_n, P_{\mathcal{S}_n})$ over S has a probability distribution equal to the joint probability distribution of $\mathcal{B}_n$, that is, $P_{\mathcal{S}_n} = P_{\mathcal{B}_n}$.*

Given the structure of the Bayesian network from Figure 1, we have $\Pi_{A_n} = \emptyset$, $\Pi_{B_n} = \{A_n\}$ and $\Pi_C = \{A_n\}$, and by Equation (1) we have:

$$P_{\mathcal{S}_n} = P(A_n)P(B_n \mid A_n)P(C \mid A_n). \tag{10}$$

where each (conditional) probability is obtained by the observed frequencies given by Equation (2). From Equation (7) and for each value of n, the difference of entropy applied to these context is represented as:

$$I(S) - I(\mathcal{S}_n) = H(\mathcal{S}_n) - H(S). \tag{11}$$

As stated above, we want to measure the loss of information about class when we lose correlations. This can be performed using conditional entropy, namely

$$H_{P_{\mathcal{S}_n}}(C \mid X_1 \ldots X_L) - H_{P_S}(C \mid X_1 \ldots X_L) = H(C \mid A_n) - H(C \mid A, B), \tag{12}$$

where $A, B = A_n, B_n$ (note that $A_n, B_n = X_1 \ldots X_L$ for all n). The conditional entropy is used to quantify the uncertainty about the classes of the collection of TS, given that T is fully or partially known. On the one hand, $H(C \mid A_n)$ consists of the amount of information required to predict the class labels, provided that the TS are known until the TP n. On the other hand, $H(C \mid A, B)$ corresponds to the amount of information needed to describe C, based on the knowledge of the entire T. The difference between these two conditional entropies measures the knowledge that the whole TS provides about the classes (correlation between C and A, B), which is not represented by the incomplete data (correlation between C and A_n). Thus, Equation (11) can be viewed as the lack of information caused by describing the structural relation C from A_n, i.e., the loss of knowledge for using the collection of TS only until the early TP in the classification process.

In addition to earliness in predicting the classes, the goal consists of finding the value n for which $\mathcal{S}_n$ represents the system S with reasonable complexity. Since this can be seen as a problem of learning the BN from Figure 1, both MDL and AIC scores are applied to the multivariate correlations for EC approach, in the interest of finding the best fitting model. These scores are used as two criteria for choosing the early TP, such that the selection of the model takes its simplicity into consideration. From Equation (3) and considering $P_{\mathcal{S}_n}$, described in Equation (10), the MDL score is defined as

$$\begin{aligned} MDL(D \mid \mathcal{S}_n) &= \frac{\log_2 N}{2} |\mathcal{S}_n \mid - \sum_{i=1}^{N} \log_2 \left[P_{\mathcal{S}_n}(A_n, B_n, C)\right] \\ &= \frac{\log_2 N}{2} |\mathcal{S}_n \mid - \sum_{i=1}^{N} \log_2 \left[P(C \mid A_n) P(B_n \mid A_n) P(A_n)\right], \end{aligned} \tag{13}$$

where N is the number of instances in the dataset D, $|\mathcal{S}_n|$ denotes the number of independent parameters in the model, and $P_{\mathcal{S}_n}$ is the underlying distribution associated to the structure $\mathcal{S}_n$, which describes S as a representation of the given data. Similarly, the AIC score, applied to this context, is defined as:

$$\begin{aligned} AIC(D \mid \mathcal{S}_n) &= |\mathcal{S}_n \mid - \sum_{i=1}^{N} \log_2 \left[P_{\mathcal{S}_n}(A_n, B_n, C)\right] \\ &= |\mathcal{S}_n \mid - \sum_{i=1}^{N} \log_2 \left[P(C \mid A_n) P(B_n \mid A_n) P(A_n)\right]. \end{aligned} \tag{14}$$

As represented in the direct acyclic graph structure from Figure 1, the goal is to analyze how the structural relation A_n is able to describe C, while the correlation between B_n and C is neglected. For this reason, the computation of the network complexity only considers the relation between the early states and the class labels,

$$|\mathcal{S}_n| = |\{A_n, C\}| = ||A_n|| - 1 + (||C|| - 1)\, ||A_n|| = ||A_n|| \times ||C|| - 1, \tag{15}$$

where $||A_n||$ and $||C||$ denote the number of distinct observations in the structural relation A_n and C, respectively. In Equations (13) and (14), the first term quantifies the complexity of the model, i.e., the amount of information required to encode not only $\mathcal{S}_n$, but also the data given $\mathcal{S}_n$. The second term measures the LL of the data based on the model, i.e., the amount of information needed to represent the dataset D according to the probability distribution $P_{\mathcal{S}_n}$. As n increases, the size of A_n becomes larger, the number of correlations is higher and, consequently, the complexity of the model increases. In addition, the more information about the TS there is, the better the correlations describe the data, which means a decrease in the number of bits needed to describe C from A_n. The difference between these two terms describes the trade-off between the model complexity and the effectiveness of the fit to the data. The simplest model, that is able to use the least amount of correlations while maintaining a distribution as close to the original as possible, is found through minimizing both $MDL(D|\mathcal{S}_n)$ and $AIC(D|\mathcal{S}_n)$.

The Multivariate Correlations for Early Classification (MCEC) procedure, summarized in Algorithm 1, receives as input a comma-separated values (CSV) file, containing the TS and the respective class labels, and a scoring function ϕ. Both univariate and multivariate TS are allowed; however, the TS must be of fixed length. For both AIC and MDL, the overall time complexity of the MCEC algorithm is $O(L^3 N^2 \log^2(N))$, where L is the size of the time-series and N is the number of time-series in the data. Indeed, the procedure needs to store and count at most N configurations of A_n, and checking whether a configuration of A_n (with size $O(L)$) already occurred takes $O(L \log(N))$ time (using a binary search tree such as an AVL tree). Moreover, for each configuration of A_n the procedure

needs to store and count at most N configurations of B_n (to establish the frequencies $P(B_n|A_n)$), which leads to an overall time complexity of $O(L^2N^2\log^2(N))$ for Step 3.

Algorithm 1 MCEC algorithm

1: **for** $n \in \{1, \ldots, L\}$ **do**
2: Compute $\mathcal{S}_n = \{\{A_n, B_n\}, \{A_n, C\}\}$
3: Compute $\phi(D \mid \mathcal{S}_n)$
4: If $\phi(D \mid \mathcal{S}_n)$ is minimum, store n
5: Output stored n

4. Experimental Results

The proposed algorithm was implemented in Java language, and is freely available in GitHub at https://joaopbeirao.github.io/MCECalgorithm/. In the interest of verifying the reliability of the proposed EC approach, an investigation on the accuracy of multiple classifiers was done. In this regard, we vary the length of the TS and check if the proposed cut-points attain a similar accuracy when compared with the complete TS. We note that the purpose of this assessment is not to elicit the best classifier among those used, but rather to understand the consequences of using a truncated time-series in classification tasks. Seven classifiers were considered, using the default parameters and stratified cross-validation with 10 folds: Naïve Bayes (NB), Bayes Net (BN), Sequential Minimal Optimization (SMO), C4.5 decision tree (J48), Reduces Error Pruning Tree (REPTree), Forest of Multiple Random Trees (RandomForest) and *k*-Nearest-Neighbor (*k*NN). We performed classification using the classifiers from the Weka Data Mining Software (version number 3.8) [25]. All the experiments were conducted using a PC with an Intel Core i7-2677M@1.80 GHz CPU and with 4 GB RAM.

From the MCEC algorithm, for each dataset, three values for the EC TP (n) were extracted. The first value is obtained from the difference in entropy measure: the smallest n, such that $H(C \mid A_n) - H(C \mid A, B) \leq 0.3 \times [H(C \mid A_1) - H(C \mid A, B)]$, which means that n corresponds to the TP where a reduction of 70% from the initial value of entropy is verified, henceforth called $CH-70$, defined as the score:

$$CH-70(D|\mathcal{S}_n) = \begin{cases} +\infty & \text{if } H(C|A_n) - H(C|A,B) > 0.7(H(C|A_1) - H(C|A,B)) \\ n & \text{otherwise.} \end{cases}$$

The second and third values are the result of minimizing $MDL(D \mid \mathcal{S}_n)$ and $AIC(D \mid \mathcal{S}_n)$, respectively, i.e., n consists of the TP where the scores are minimum.

For analyzing the performance of the MCEC algorithm, we use thirteen benchmark datasets from the UEA & UCR Time Series Classification Repository [26,27]. This subset of examples is considered representative, as it comprises a diverse range of both dimensional parameters and classification conditions. Each dataset is composed of numeric TS with a fixed length, and their respective class labels.

For each example, a training set and a test set are provided separately. The preprocessing of the data included the aggregation of both training and test subsets in one single dataset. In addition, a TS discretization was performed following the guidelines proposed by Lin et al. [28] and Hu et al. [29]; no feature selection was performed. None of the datasets contained missing values; therefore, no imputation was required. The description of the data used in the experiments is given in Table 1. For each dataset, the number of class values (#C), the number of variables for each TP (m), the length of the discretized TS (L), the number of instances (N), and the type of data are provided.

Table 1. Description of the benchmark time-series used in the experiments.

Dataset	#C	m	L	N	Type
ArrowHead	3	1	14	211	image
BirdChicken	2	1	25	40	image
Car	4	1	25	120	sensor
CBF	3	1	12	930	simulated
Coffee	2	1	71	56	spectro
Computers	2	1	5	500	device
ECG200	2	1	10	200	EGC
GunPoint	2	1	150	200	motion
Meat	3	1	7	120	spectro
SynthControl	6	1	6	8	simulated
Wafer	2	1	152	7164	sensor
ECG200M	2	2	13	200	ECG
WaferM	2	6	13	1195	sensor

The results from Table 2 describe the MCEC algorithm effort in attempting EC, based on the analysis of the information contained in the datasets. For each dataset, it is given the Earliness as the percentage value associated with the EC TP n computed from MCEC algorithm: Earliness$[\%] = \frac{n}{L} \times 100$. The Accuracy columns present only the best (10-fold cross-validation) classification result (among all classifiers used) for the given data, predicting the class variable at TP n. The column Full contains the best outcome (among all classifiers used) for the complete TS, and it is used as a reference framework.

Table 2. Experimental results of the MCEC algorithm.

Dataset	Earliness			Accuracy			Full
	$CH-70$	MDL	AIC	$CH-70$	MDL	AIC	
ArrowHead	57.14%	7.14%	42.86%	72.51%	38.39%	70.14%	78.20%
BirdChicken	20.00%	4.00%	8.00%	62.50%	57.50%	75.00%	77.50%
Car	52.00%	12.00%	40.00%	70.83%	40.83%	70.83%	65.83%
CBF	66.67%	33.33%	66.67%	91.61%	74.09%	91.61%	94.95%
Coffee	19.72%	9.86%	9.86%	85.71%	76.79%	76.79%	98.21%
Computers	20%	20%	80%	51.20%	51.20%	60.80%	61.40%
ECG200	50.00%	20.00%	30.00%	87.50%	76.50%	87.50%	85.50%
GunPoint	24.00%	0.67%	15.33%	92.00%	71.50%	83.00%	99.50%
Meat	28.57%	28.57%	57.14%	69.00%	69.00%	74.50%	75.50%
SynthControl	62.50%	25.00%	50.00%	89.50%	62.67%	85.00%	92.33%
Wafer	7.24%	1.32%	1.97%	97.91%	97.60%	97.66%	99.85%
ECG200M	38.46%	7.69%	15.38%	80.50%	73.00%	80.00%	81.00%
WaferM	30.77%	7.69%	30.77%	93.22%	89.36%	93.22%	96.15%

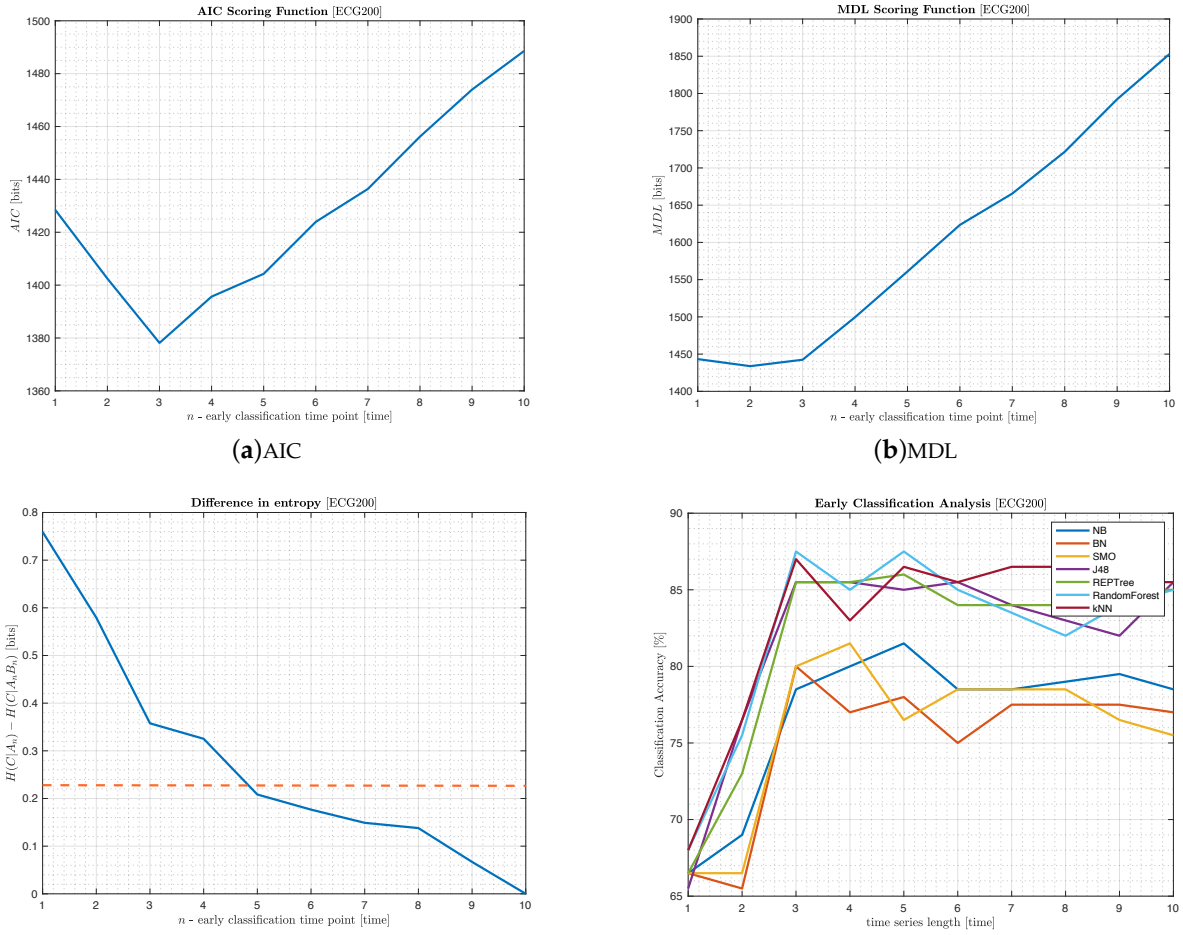
The results from the Earliness column of Table 2 confirm that in most datasets, it is viable to perform early classification. Concerning the $CH-70$ and AIC, the classification accuracy with fewer TPs outperforms the reference value (Full column) for the Car and ECG200 dataset. These examples suggest that it is possible to obtain a better classification performance using only part of the TS (in these cases, around 50% or even less) from the data. In the case of the Computers dataset, AIC indicates the use of 80% of the TS, attaining with it an accuracy very close to the reference value. A similar result is obtained with the Meat dataset, where AIC pointed out to use just over half the size of the TS. In the multivariate ECG200M dataset, AIC pinpoint the use of only 15% of the MTS obtaining a difference of 1% less in accuracy, when compared with the reference value in the Full column. In other datasets (ArrowHead, BirdChicken, SynthControl, Wafer, and WaferM) accuracy outcomes with at most -3% in comparison with the full-length result are obtained. This means that, in these experiments, with fewer TPs analyzed (earlier in time), the loss in terms of classification accuracy can be diminished.

Regarding the scoring metrics, MDL always proposes the lowest values for the early classification TP (Earliness column). However, the corresponding classification accuracy results of $CH-70$ and AIC outperform the ones for MDL in all cases. Moreover, $CH-70$ achieves higher classification accuracy in six out of thirteen cases, AIC in three datasets, and in the remaining four instances, a draw is verified between the $CH-70$ and AIC. Therefore, from the experimental tests described in Table 2, in general, $CH-70$ achieves better results, in terms of classification accuracy, and MDL demonstrates a superior earliness ability. AIC evidences the foremost competence in balancing these two targets. Nevertheless, the early classification capabilities of the MCEC algorithm are acknowledged, seeing that this context is based on the trade-off between these two main objectives: accuracy and earliness.

4.1. Analysis of the ECG200 and Car Results

Herein, we detail the analysis of two datasets that illustrate the merits of the proposed method, complementing Table 2 with a graphical interpretation of the results, drawing conclusions both in terms of the relevant fraction of the time-series data for classification and the quality of the proposed scoring criteria.

The ECG200 data comprises two hundred time-series that traces the electrical activity recorded during one heartbeat used as part of R. Olszewski Ph.D. thesis [30]. It aims to discriminate between normal heartbeat and myocardial infarction. The original time-series has 96 TPs. After discretization, following the guidelines proposed by Hu et al. [29], the time-series was reduced to ten TPs. The results are depicted in Figure 2.



(**a**)AIC (**b**)MDL

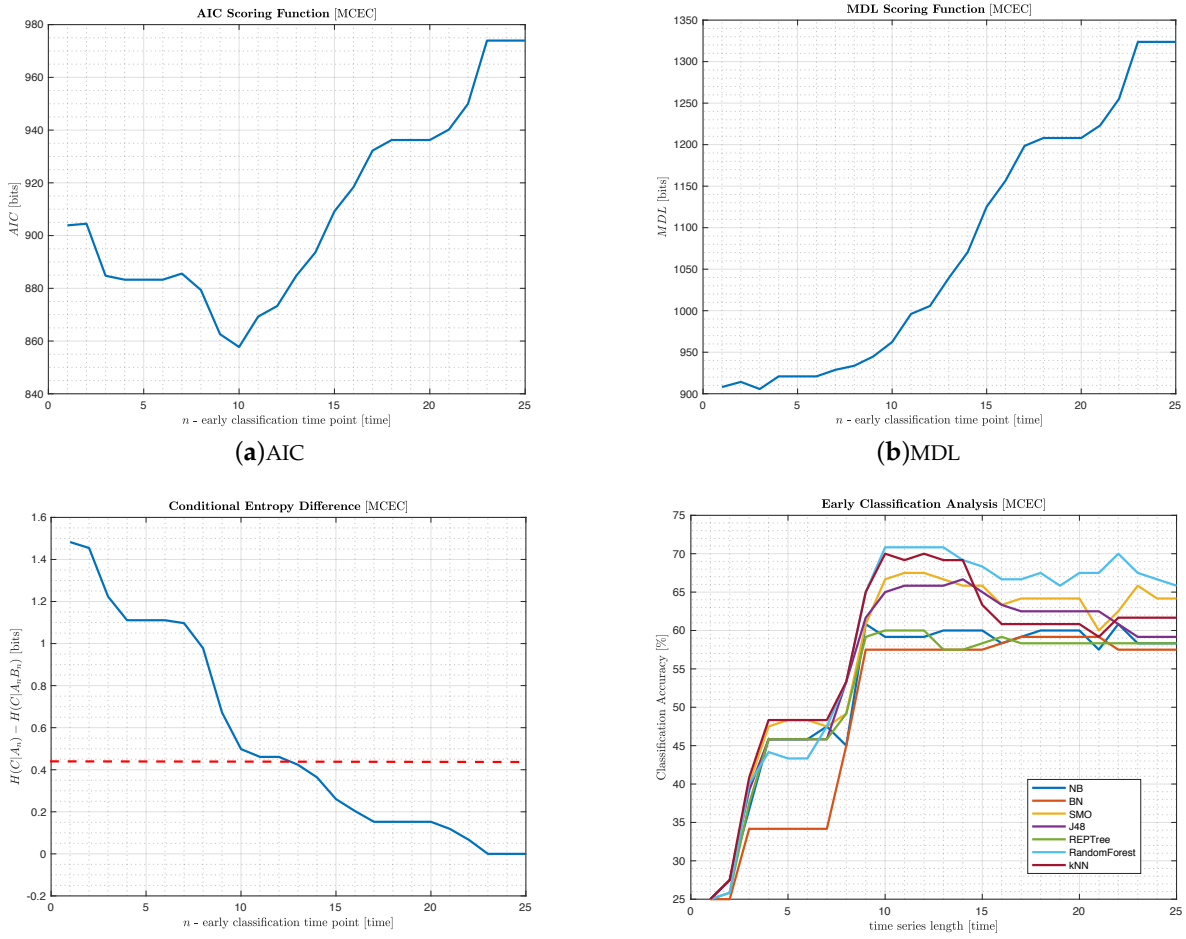
(**c**)Difference in entropy (**d**)Classification results

Figure 2. Experimental results of the MCEC algorithm on the ECG200 dataset.

Results suggest that only a fraction of the heartbeat signal is crucial to discriminate between classes. In particular, AIC (Figure 2a) and $CH-70$ (Figure 2c) proposed to cut the time-series at TPs 3 and 5, respectively; in $CH-70$ the cut point corresponds to the first n bellow the dashed red line. Indeed, both cut points proposed by AIC and $CH-70$ agree upon a peak in accuracy of 87.5% (c.f. Table 2), and so surpassing the full-length time-series in 2%. The MDL criterion is more confident in this regard (c.f. Figure 2b), proposing to cut the time-series at $n=2$. However, in this particular case, it suggested a premature cut leading to lower prediction values, as depicted in Figure 2d.

In addition, the experimental results on the Car example are represented in Figure 3. Figure 3a,b represent the variation of $AIC(D \mid \mathscr{S}_n)$ and $MDL(D \mid \mathscr{S}_n)$, respectively, for $n \in \{1,\ldots,25\}$. While for AIC, the minimum is reached at $n=10$ (corresponding to 40% of 25 TPs), for MDL, the lowest value is attained at $n=3$ (corresponding to 12% of 25 TPs). In both cases, this extreme is followed by an irregular growth until $n=23$, where it stabilizes at a maximum value. Figure 3c describes the behaviour of $H(C \mid A_n) - H(C \mid A,B)$ while varying n from 1 to $L=25$. A decrease of 70% from the initial entropy value is obtained at $n=13$ (corresponding to 52% of 25 TPs), depicted with a dashed red line.

Figure 3d includes the classification accuracy of the Car dataset. Note that there are two jumps in classification accuracy at $n=3$ and $n=10$, corresponding precisely to the timestamps elicited by the MDL and AIC, respectively.



(**a**)AIC (**b**)MDL

(**c**)Difference in entropy (**d**)Classification results

Figure 3. Experimental results of the MCEC algorithm on the Car dataset.

4.2. Statistical Significance

Experimental results were compared with statistical significance tests in order to understand the benefit of the trade-off between the two main goals in EC: accuracy and earliness. Among the tested datasets, the MCEC algorithm provided a value of n, with an associated percentage (Earliness). For each situation, the group of classifiers determined the Accuracy value. In addition, the classification of the full-data worked as a reference framework: no earliness and complete TS accuracy. In order to represent the balance between these two requirements, we consider the following quantity $BEA(p) = p \times (100 - E) + (1 - p) \times A$, where E and A correspond to the Earliness and Accuracy percentages, respectively, and p consists of the weight that determines the relevance given to each variable. Seeing that an accurate classification is desirable, as early as possible, BEA describes the management of the two fundamental challenges of the EC problem. The thirteen datasets from Table 2 were considered, as well as their respective values of E and A, for each of the three measures that compose the MCEC algorithm, together with the reference framework. Note that all Full outcomes verify $E = 0\%$, since the entire collection of TS is considered for classification.

Table 3 includes the results of the Wilcoxon signed-rank sum test [31], for comparing the performance of classification using the MCEC algorithm timestamp n on the TS. These tests examine the relation between the scores in pairs, in order to verify if there is enough evidence to claim that the differences are significant, for a significance level of $\alpha = 0.05$. The arrow in Table 3 points towards the measure with better performance, according to the value of $p \in \{0, 0.25, 0.5, 0.75\}$. The double arrow means there is enough evidence to claim the difference is significant.

Table 3. Comparison of the classification using the three MCEC algorithm cuts ($CH - 70$, AIC and MDL) and Full TS against each other, using the Wilcoxon signed-rank sum test applied to the trade-off experimental data, scored according to $BEA(p)$.

Comparison		***p***			
		0	**0.25**	**0.5**	**0.75**
$CH-70 \Leftrightarrow MDL$	size	26	26	26	26
	p-value	<0.01	<0.01	0.81	<0.01
	better	$\Leftarrow$	$\Leftarrow$	$\rightarrow$	$\Rightarrow$
$CH-70 \Leftrightarrow AIC$	size	25	26	26	26
	p-value	<0.01	0.01	0.34	<0.01
	better	$\Leftarrow$	$\Leftarrow$	$\rightarrow$	$\Rightarrow$
$CH-70 \Leftrightarrow$ Full	size	26	26	26	26
	p-value	<0.01	<0.01	<0.01	<0.01
	better	$\Rightarrow$	$\Leftarrow$	$\Leftarrow$	$\Leftarrow$
$MDL \Leftrightarrow AIC$	size	20	20	20	20
	p-value	<0.01	<0.01	0.23	0.37
	better	$\Rightarrow$	$\Rightarrow$	$\rightarrow$	$\leftarrow$
$MDL \Leftrightarrow$ Full	size	26	26	26	26
	p-value	<0.01	0.70	<0.01	<0.01
	better	$\Rightarrow$	$\rightarrow$	$\Leftarrow$	$\Leftarrow$
$AIC \Leftrightarrow$ Full	size	26	26	26	26
	p-value	<0.01	0.37	<0.01	<0.01
	better	$\Rightarrow$	$\leftarrow$	$\Leftarrow$	$\Leftarrow$

The results demonstrate that, for $p = 0$, there is enough evidence to claim that Full surpasses all other measures. Furthermore, between $CH - 70$ and the other two model selection criteria, the difference in entropy outperforms both scoring functions, and AIC shows better results than MDL. All these differences are statistically significant. For $p = 0.25$, $CH - 70$ has the best performance in comparison with all of the remaining. The AIC measure seems to achieve significantly superior results than MDL; however, there is not enough evidence to claim that AIC outperforms Full, nor that the latter surpasses MDL. For $p = 0.5$, the only assurance consists of Full performing the worst. Among $CH - 70$, MDL, and AIC, the differences between them are not statistically significant. Lastly, at $p = 0.75$, Full continues to be surpassed by all the others, as well as the difference in entropy in

comparison with both model selection criteria. However, between MDL and AIC, there is not enough evidence to confirm which performs the best.

5. Conclusions

This work proposes a novel algorithm, named MCEC, that aids in addressing the challenges associated with the task of early prediction in (univariate and multivariate) time-series data. Existing methods yield, for each time-series sample, a timestamp from which it is possible to perform early classification, failing to provide an overall data perspective. To the best of our knowledge, MCEC is the first approach that can grant the early opportunity of the entire data, allowing us to reason about prediction outcomes after understanding data idiosyncrasies.

MCEC is very flexible as it can be used with different scoring criteria, allowing for the trade-off between earliness and accuracy. We propose three measures: $CH-70$, MDL and AIC. The achieved results confirm the ability of the MCEC method to examine the EC opportunity within a dataset. In general, the three criteria are capable of choosing a timestamp for which the time-series classification is plausible. Overall, the $CH-70$ obtains better accuracy results, MDL demonstrates a superior tendency for earliness, and AIC attains the most competent balance between both aims. Examples, where the earliness is very low, may indicate that, given the information available, the criterion recognized that the increase in the knowledge obtained from the data did not justify the growth in the model complexity required for its description. Conversely, the AIC results demonstrate a more adventurous disposition in choosing the value for the early classification timestamp, which produced a relative success in benchmark data.

In terms of future work, several machine learning tasks can be developed based on the capabilities of this information-theoretic approach. A classification method can take profit from it, giving different attention to time-series whose classification timestamp deviates from that derived from the data; for instance, inspecting only part of electrocardiograph signals might improve existing classification methods [32]. It can also aid in determining the change point detection, as well as feature extraction and selection methods from multivariate time-series data. In the latter case, for instance, a greedy feature selection could be performed based not only on the difference in entropy measure but also on the model selection criteria. Finally, MCEC, if applied in a time-series from back/present to front/past, has the potential of unraveling the optimal Markov length of the stochastic process underlying the data.

Author Contributions: M.L. performed the theoretical derivations and formal analysis. J.P.B. implemented the software and wrote the first draft of the manuscript. N.P., A.M.C. and P.M. were responsible for the conceptualization of the work, supervised its implementation, and suggested the methodology. A.M.C. was responsible for funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by national funds through FCT, Fundação para a Ciência e a Tecnologia, under contract IT (UID/EEA/50008/2019), projects PREDICT (PTDC/CCI-CIF/29877/2017), PERSEIDS (PTDC/EMS-SIS/0642/2014), NEUROCLINOMICS2 (PTDC/EEI-SII/1937/2014), and QuantumMining (POCI-01-0145-FEDER-031826).

Acknowledgments: M.L. acknowledges FCT PhD grant PD/BD/114334/2016 and the DP-PMI FCT programme. N.P. acknowledges FCT Scientific Employment Stimulus grant CEECIND/04594/2017/CP1393/CT0006.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Theorem 1

The proof is based on the Lagrange multiplier method. For the sake of simplicity, take $A = A_n$ and $B = B_n$. Denote

$$p_{abc} = \hat{P}_S(A = a, B = b, C = c)$$

and

$$q_{abc} = \hat{P}_S(B = b \mid A = a)\hat{P}_S(C = c \mid A = a)\hat{P}_S(A = a) = P_{\mathcal{B}_n}(A = a, B = b, C = c).$$

We proceed to show that q_{abc} is a stationary point of the entropy function

$$H(q_{abc}) = -\sum_{a,b,c} q_{abc} \log(q_{abc})$$

under the constrains

- $p_{ab} = \sum_c p_{abc} = \sum_c q_{abc} = q_{ab}$,
- $p_{ac} = \sum_b p_{abc} = \sum_b q_{abc} = q_{ac}$, and
- $\sum_{a,b,c} q_{abc} = 1$.

By the concavity of entropy, any solution to the previous problem is a maximum. Using the Lagrange multiplier method, we look for a maximum of the function

$$\mathcal{L}(q'_{abc}) = H(q'_{abc}) - \sum_{a,b} \lambda_{ab}(q'_{ab} - p_{ab}) - \sum_{a,c} \lambda_{ac}(q'_{ac} - p_{ac}) - \lambda_u(\sum_{a,b,c} q'_{abc} - 1),$$

where the $\lambda_{ab}, \lambda_{ac}$ and λ_u are Lagrange multipliers. It is now enough to show that by setting $q'_{abc} = q_{abc}$, and for adequate values of the Lagrange multipliers, $\frac{\partial \mathcal{L}}{\partial q'_{abc}}$, $\frac{\partial \mathcal{L}}{\partial \lambda_{ab}}$, $\frac{\partial \mathcal{L}}{\partial \lambda_{ac}}$, $\frac{\partial \mathcal{L}}{\partial \lambda_u}$ all vanish. Effectively,

$$\begin{aligned}
\left.\frac{\partial \mathcal{L}}{\partial q'_{abc}}\right|_{q_{abc}} &= -\left.(\log(q'_{abc}) + 1 + \lambda_{ab} + \lambda_{ac} + \lambda_u)\right|_{q_{abc}} \\
&= -(\log(p_{ab}p_{ac}/p_a) + 1 + \lambda_{ab} + \lambda_{ac} + \lambda_u) \\
&= \log(p_a) - \log(p_{ab}) - \log(p_{ac}) - 1 - \lambda_{ab} - \lambda_{ac} - \lambda_u
\end{aligned}$$

vanishes by setting

$$\lambda_{ab} = \frac{1}{2}\log(p_a) - \log(p_{ab}), \quad \lambda_{ac} = \frac{1}{2}\log(p_a) - \log(p_{ac}), \quad \lambda_u = -1.$$

On the other hand, we have that

$$\begin{aligned}
\left.\frac{\partial \mathcal{L}}{\partial \lambda_{ab}}\right|_{q_{abc}} &= \left.(q'_{ab} - p_{ab})\right|_{q_{abc}} = \sum_c q_{abc} - p_{ab} = \sum_c \frac{p_{ab}p_{ac}}{p_a} - p_{ab} = p_{ab} - p_{ab} = 0 \\
\left.\frac{\partial \mathcal{L}}{\partial \lambda_{ac}}\right|_{q_{abc}} &= \left.(q'_{ac} - p_{ac})\right|_{q_{abc}} = \sum_b q_{abc} - p_{ac} = \sum_b \frac{p_{ab}p_{ac}}{p_a} - p_{ac} = p_{ac} - p_{ac} = 0.
\end{aligned}$$

Finally,

$$\left.\frac{\partial \mathcal{L}}{\partial \lambda_u}\right|_{q_{abc}} = \left.\sum_{a,b,c} q'_{abc} - 1\right|_{q_{abc}} = \sum_{a,b,c} q_{abc} - 1 = 0,$$

as q_{abc} is a probability distribution by construction. Therefore, $q_{abc} = P_{\mathcal{B}_n}$ satisfies the definition of the probability distribution associated with the structure $\mathcal{S}_n = \{\{A, B\}\{A, C\}\}$, and therefore, it matches $P_{\mathcal{S}_n}$.

Appendix B. Results on Synthetic Data

Herein we describe the empirical study of the proposed method on synthetically generated datasets. The procedure to generate synthetic data is based on the exclusive disjunction, and it allows an interpretation of the results in comparison with the expected outcomes. The parametrization of the data generator enables the variation on standard time-series dataset aspects: the number of features per time point (m), the length of the time-series (L), and the number of instances (N). Moreover, two

additional variables are included: the number of randomly generated columns (x) and the percentage of noise in the dataset ($pNoise$). Recall the data type is boolean for all features, and all datasets contain 2 classes. According to the specified parameters, a database is created, with N time-series, each with m attributes per time point and length equal to L. The value of x represents the number of initial instants that are randomly generated. The following time points are computed as the XOR of the x previous ones. For the multivariate case, where each instant is composed of a set of features, the process is maintained for each attribute independently. The class labels are computed with the same use of the exclusive disjunction, however, for $m \geq 2$, another XOR is applied to the collection of features, in order to obtain only one value for the class attribute. Aiming for providing more realistic data, the noise is used, causing a number of arbitrary positions to be switched, i.e., 0 becomes 1 and vice versa.

The idea is to produce a set of time-series where the class labels are a function of the x initial time points in the interest of analyzing if the proposed algorithm can recognize this correlation and, consequently, the early classification opportunity.

At the first stage, the impact of the dimensional parameter variation on the MCEC method behavior is analyzed. Seeing that the two model selection criteria used in the proposed approach are sensitive to the data size, the variation of the number of variables m is studied under different conditions. Therefore, the output graphs are examined except for the classification accuracy, since the intention is to explore how the system is affected by modifications in the size of the dataset. The absolute values of the log-likelihood and of both scoring functions increase with the number of variables m. Because of that, feature scaling normalization, given by:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{A1}$$

is applied to the results, for comparing the relative behaviour of the quantities. Only the entropy graph includes the non-normalized values, on account of being a difference between two variables of the same order of magnitude.

Figure A1 represents the behavior of the four measures under investigation for datasets with a different number of instances.

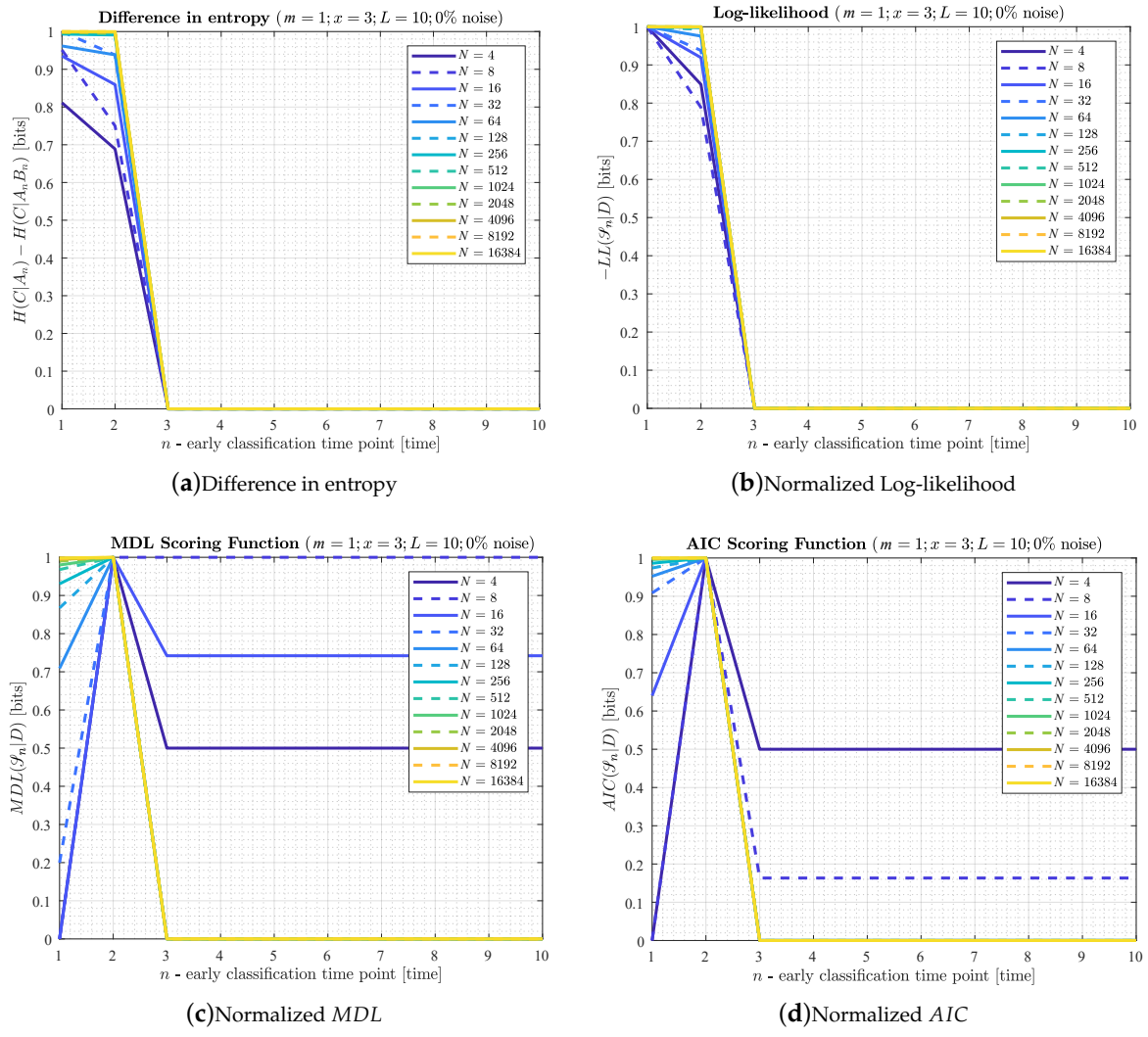
As previously mentioned, $H(C \mid A_n) - H(C \mid A, B)$ quantifies the lack of knowledge caused by describing the classes using the time-series in the dataset only until time point n. From Figure A1a, the variation of m does not extensively affect the difference in entropy. Since $H(C \mid A_n) - H(C \mid A, B) = 0$ for $n \geq 3$, there seems to be enough information to predict the class labels, with the first three-time points. The variation of entropy from $n = 1$ to $n = 2$ is sharper for lower values of N, and it becomes null while the number of instances increases.

Furthermore, $-LL(D \mid \mathcal{S}_n)$ describes the amount of information needed to represent the dataset D using the model $\mathcal{S}_n$. Figure A1b demonstrates that the data is entirely depicted by the structure $\mathcal{S}_3$, seeing that the normalized log-likelihood is zero from $n = 3$ forward. The behavior of this measure is very similar to the difference in entropy since they both quantify how good the model fits the data.

Regarding the scoring functions, in the early classification context, the lowest value of both scores corresponds to the time point from which additional information can be disregarded. While the graph from Figure A1c shows that MDL has a minimum at $n \geq 3$ for $N \geq 32$, the one from Figure A1d displays AIC achieving it for $N \geq 16$. In the two cases, the scores are constant from the point where they attain the lowest value. This is because each time point is a function of what is behind since it consists of the exclusive disjunction of the three previous instants. Consequently, the number of independent parameters in group A_n ($||A_n||$) is constant for $n \geq 3$, i.e., the number of distinct cases in the list $\mathcal{A}_n$ stabilises from that point on.

Three columns are randomly generated ($x = 3$), which means that a feasible prediction of the class labels is expected using only the first three-time points. This figure describes the univariate case ($m = 1$), for a fixed time-series length ($L = 10$), with no addition of noise ($pNoise = 0\%$). In order

to explore a dimensional range for the data size, the values for m comprise a set of powers of two: $\{2^2, 2^3, \ldots, 2^{14}\}$.



(**a**)Difference in entropy

(**b**)Normalized Log-likelihood

(**c**)Normalized *MDL*

(**d**)Normalized *AIC*

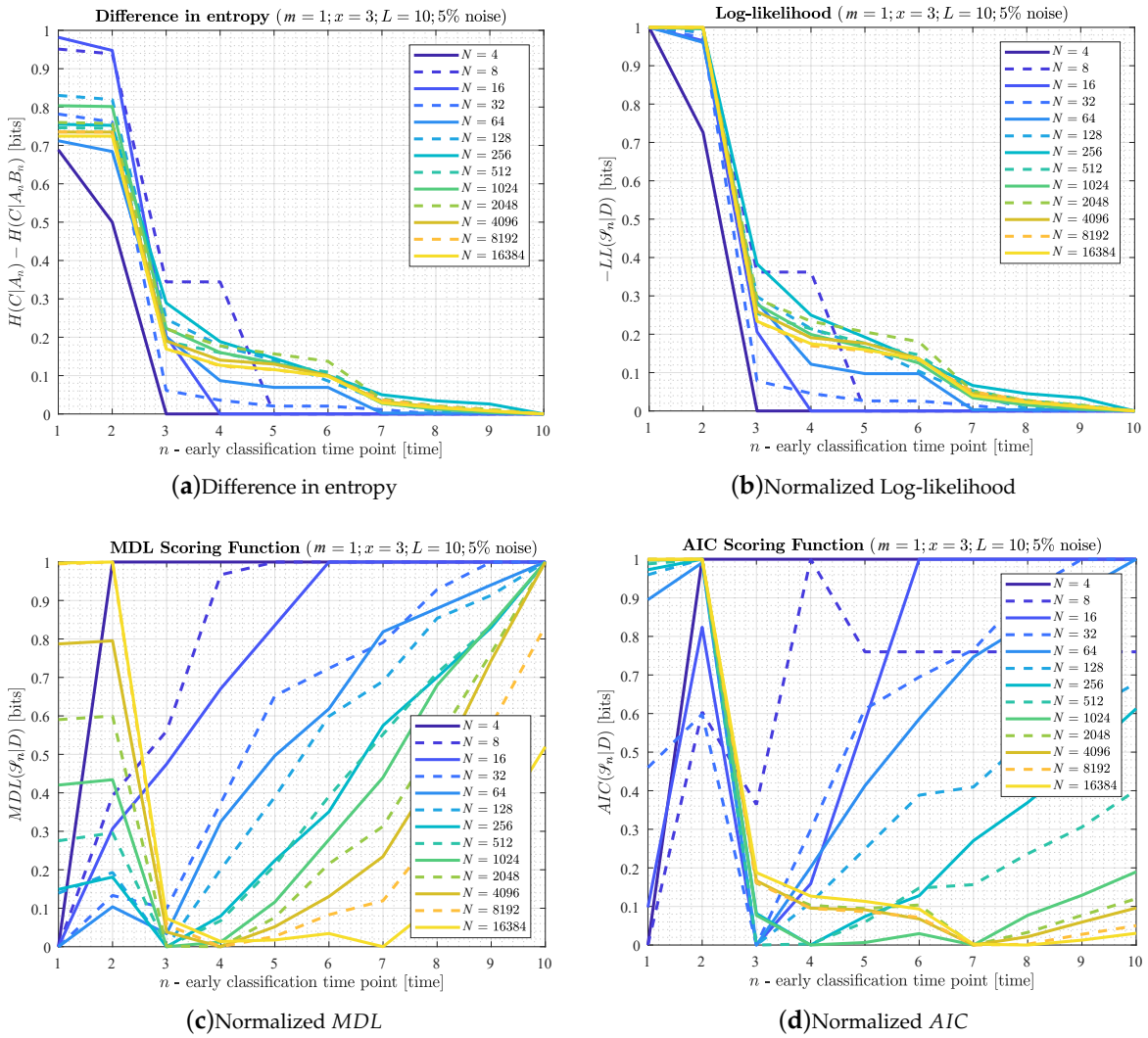
Figure A1. Variation of n for synthetic datasets with $m = 1$, $L = 10$, $x = 3$, $pNoise = 0\%$.

Figure A2 describes the experimental tests in datasets with the same parameters as the ones used in Figure A1, except for the percentage of noise.

With $pNoise = 5\%$, a more realistic environment is simulated. The difference in entropy (Figure A2a) and the log-likelihood (Figure A2b) have a smoother decreasing behavior and more difficulty in reaching zero, in particular for higher values of m. However, in general, the most significant reduction in both cases is verified from $n = 2$ to $n = 3$. This indicates that, although the lack of information is minimized as more of the time-series is observed, for a certain threshold, the graphs show an early classification opportunity. Similarly to the 0% noise case, the increase in the number of instances is followed by a stabilization from $n = 1$ to $n = 2$, which means that less knowledge is gained with the use of only the first two time points. The behavior of both measures seems to be convergent for $N \to \infty$.

Regarding the scoring functions, Figure A2c and Figure A2d show a high variance of *MDL* and *AIC* with the data size, respectively. For few instances, the lowest value is obtained at $n = 1$, which means that the model, considered the best in terms of complexity and fitness to the data, is the one with merely the first time point. In this case, a proper model selection is impracticable, since the samples available are insufficient to extract the model conveniently; that is, the data does not contain

enough information. MDL displays a minimum at $n \geq 3$ for $N \geq 128$ (higher than for 0% noise) and AIC for $N \geq 16$ (the same as for 0% noise). The lowest value of $MDL(D \mid \mathscr{S}_n)$ is attained at $n = 3$ for $N \in \{128, \ldots, 1024\}$, at $n = 4$ for $N \in \{2048, \ldots, 8192\}$, and at $n = 7$ for $N = 16,384$. The minimum of $AIC(D \mid \mathscr{S}_n)$ is reached at $n = 3$ for $N \in \{16, \ldots, 128\} \cup \{512\}$, at $n = 4$ for $N = 256$, at $n = 7$ for $N \in \{1024, \ldots, 8192\}$, and at $n = 8$ for $N = 16,384$. For the experimental range of sample size, the results suggest that, although the AIC score always elicits an EC cut point at $n \geq 3$, it is only equal to 3 for lower values of N. As N grows, the EC cut point also becomes larger. This is expected, as AIC penalizes less the complexity than MDL.

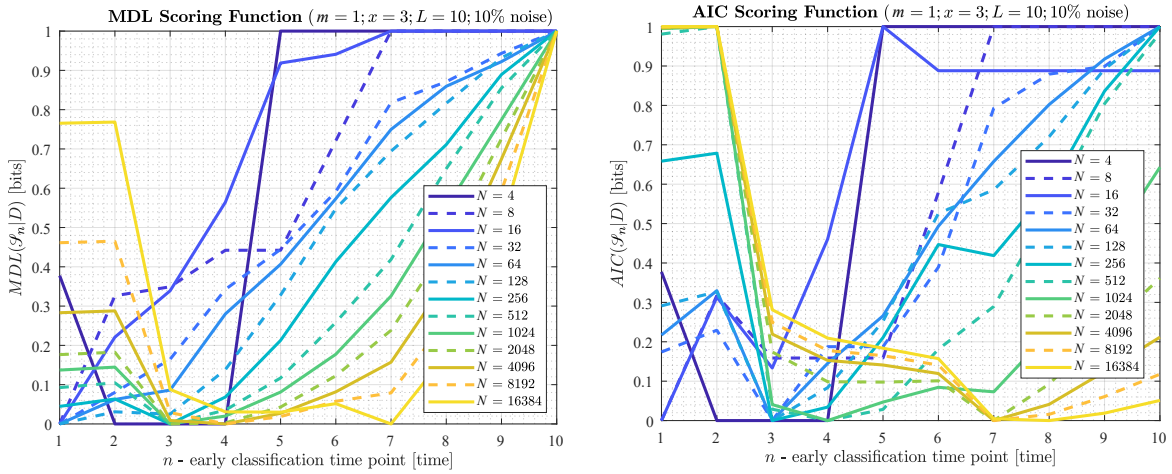


(**a**)Difference in entropy (**b**)Normalized Log-likelihood

(**c**)Normalized MDL (**d**)Normalized AIC

Figure A2. Variation of n for synthetic datasets with $m = 1$, $L = 10$, $x = 3$, $pNoise = 5\%$.

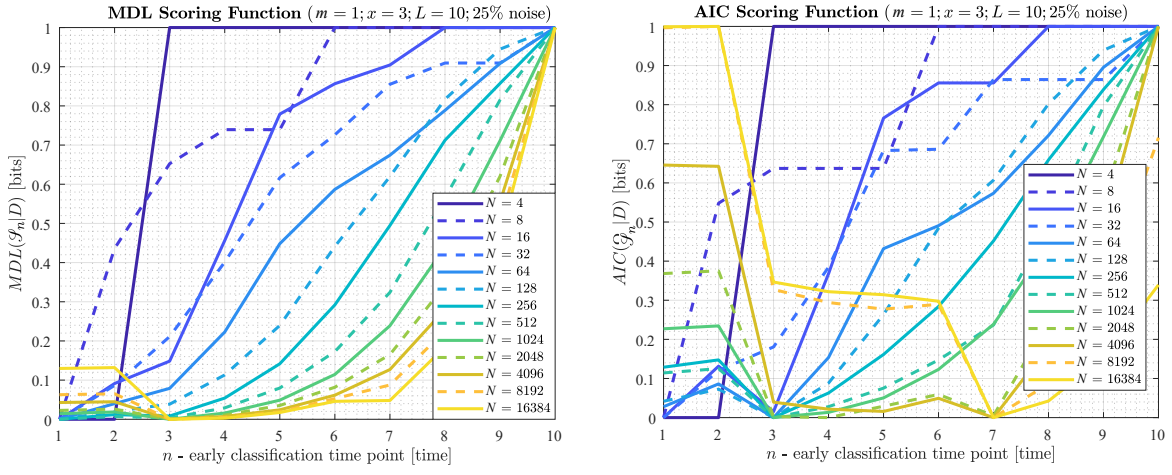
In order to examine the impact of the noise in the inferences drawn about the model selection criteria, similar experiments were performed on datasets with $pNoise$ equal to 10% and 25%. Concerning the difference in entropy and the log-likelihood measures, the decreasing behavior is preserved, although the variation becomes less accentuated with noisier data. Moreover, since noise causes uncertainty, the jump from $n = 2$ to $n = 3$ is not expressive, and consequently, the early classification opportunity at $n = 3$ is less obvious. For $pNoise = 10\%$, while the lowest value of MDL (Figure A3a) at $n \geq 3$ is attained for $N \geq 256$ (higher than for 5% noise), in AIC (Figure A3b) this minimum is reached for $N \geq 32$ (higher than for 5% noise).



(**a**)Normalized *MDL* (**b**)Normalized *AIC*

Figure A3. Variation of n for synthetic datasets with $m = 1$, $L = 10$, $x = 3$, $pNoise = 10\%$.

Note that, in Figure A3, the curve $N = 4$ also displays a minimum for $n \in \{2, \ldots, 4\}$. This event should not be considered relevant since the dataset is so reduced that noise has an unbalanced influence on the results. Proof of that is, for example, the curve $N = 8$, which does not have a minimum at $n \geq 3$. For $pNoise = 25\%$, the lowest value of MDL (Figure A4a) at $n \geq 3$ is attained for $N \geq 1024$ (higher than for 10% noise), whereas in AIC (Figure A4b) it is reached for $N \geq 64$ (higher than for 10% noise). While Figure A3a ($pNoise = 10\%$) shows the MDL graph with some ambiguity in selecting the true model for larger values of N, Figure A4a ($pNoise = 25\%$) describes the same score identifying $n = 3$ as the early time point with zero error. Furthermore, a lower deviation from the true distribution is also observed in AIC, for $pNoise = 25\%$ (Figure A4b), in comparison with the case with $pNoise = 10\%$ (Figure A3b).



(**a**)Normalized *MDL* (**b**)Normalized *AIC*

Figure A4. Variation of n for synthetic datasets with $m = 1$, $L = 10$, $x = 3$, $pNoise = 25\%$ and stuff.

A few considerations about the response of the MCEC method to variations on the size of the dataset can be referred to. Firstly, concerning the univariate context and for datasets with time-series of fixed length, the number of instances has a significant impact on both scoring functions and a not so strong influence in the difference in entropy and the log-likelihood measures. Besides, the results suggest there is a value of N from which the minimization of the model selection criteria is achieved at $n = 3$. This indicates that the number of instances in a dataset influences the effectiveness of the scoring functions in selecting the actual distribution. It is well known that both model selection criteria

are unsuitable for reduced datasets, where the number of instances is not considerably higher than the number of estimated model parameters, leading to overfitting.

Concerning the comparison between both criteria, the experiments demonstrate that, in general, AIC outperforms MDL, for more reduced datasets. However, for larger values of N, the AIC evidences a more significant deviation from the actual distribution, tending to choose more complex models than MDL. This fact verifies the MDL reputation of being more consistent than AIC in selecting the underlying model among the candidates, provided that the true model is in the set of alternatives.

In general, the sharp decreases in $H(C \mid A_n) - H(C \mid A, B)$ and in $LL(D \mid \mathcal{S}_n)$ for $n = 3$, together with the minimum values depicted in both scores, give confidence in the early classification potential of the proposed method. On the other hand, the experiments demonstrate that the decision upon the early time point (n) is not always unanimous among the three measures that compose the MCEC algorithm. This means that, in some cases, the instant from which the remaining of the time-series in the dataset can be neglected is not uniquely identified.

Additional experiments were performed to the proposed method in order to analyze the impact of the variation of two other parameters: the number of features (m) and the time-series length (L). The objective consists of not only examine the early prediction opportunity, but also continue the investigation on how the size of the dataset influences the model selection criteria.

Seeing that the algorithm is capable of handling multivariate time-series ($m \geq 2$), the study involves randomly generated datasets with $m \in \{1, 2, 3, 5\}$, while $x = 3$, $L = 10$ and $pNoise \in \{0\%, 5\%\}$. With regard to the difference in entropy and the log-likelihood, the decreasing behavior of these measures is not substantially affected by the variation on the number of features per time point. In general, the reduction within $n \in \{2, \ldots, 4\}$ is expressive, which indicates that, in this time period, there occurs a significant decrease on the amount of information needed to predict the time-series classes of the dataset.

Considering the scoring functions, the value of N from which both criteria display a minimum at $n \geq 3$ increases with m. Table A1 confirms this inference by describing the variation on the number of instances for which the minimum of MDL and AIC is attained at $n \geq 3$, according to the number of features per time point.

Table A1. Values of N from which the scoring functions display a minimum at $n \geq 3$. Parameters: $x = 3$, $L = 10$, $pNoise \in \{0\%, 5\%\}$ and $m \in \{1, 2, 3, 5\}$.

m	***pNoise***	***N***	
		MDL	***AIC***
1	0%	32	16
	5%	128	16
2	0%	1024	128
	5%	2048	256
3	0%	8192	1024
	5%	32768	4096
5	0%	$> 131,072$	65,536
	5%	$> 131,072$	$> 131,072$

In fact, for all experiments, the minima were reached for $n = 3$. Moreover, AIC seems to be less dependent on m than MDL, since its values of N are always lower. This suggests that, although the dataset size impacts the effectiveness of the model selection criteria, the early classification time point is identified with reliable consistency.

Another parameter examined was the length of the time-series in the dataset. Although the proposed method requires the data to have a fixed L, this value can vary from database to database.

In order to evaluate how the variation of the time-series length affects the MCEC algorithm, several experiments were performed with $L \in \{6, 10, 18, 38, 78, 158\}$, for the following fixed parameters: $x = 3$, $m = 1$ and $pNoise \in \{0\%, 5\%\}$.

Concerning the curves from $H(C \mid A_n) - H(C \mid A, B)$ and $LL(D \mid \mathscr{S}_n)$, the impact of the variation of L is not significant. Table A2 includes the values of N from which both scoring functions show a minimum at $n \geq 3$.

Table A2. Values of N from which the scoring functions display a minimum at $n \geq 3$. Parameters: $x = 3$, $m = 1$, $pNoise \in \{0\%, 5\%\}$ and $L \in \{6, 10, 18, 38, 78, 158\}$.

L	***pNoise***	***N***	
		MDL	***AIC***
6	0%	32	16
	5%	64	32
10	0%	32	16
	5%	128	16
18	0%	32	16
	5%	128	16
38	0%	32	16
	5%	32	32
78	0%	32	16
	5%	64	32
158	0%	32	16
	5%	64	16

Unlike the results from Table A1, the lowest values of $MDL(D \mid \mathscr{S}_n)$ and $AIC(D \mid \mathscr{S}_n)$ were not consistently obtained for $n = 3$, but instead, they deviated from the true distribution ($n \in \{4, 6, 7, 8\}$) with the increase of the number of instances. The results demonstrate that the time-series length does not considerably condition the ability of both criteria to select the best model since the values of N in Table A2 do not significantly change with the variation of L. Although not always according to the expected model ($n = 3$), and occasionally in a non-unanimous decision situation, the early classification opportunity is observable in the majority of the cases. One way of understanding this phenomenon is to notice that both the criteria are trying to model the noise itself by adding extra data points, which allows for correcting the noise.

In sum, these are the conclusions that can be drawn from the performed experiments based on the variation of the dataset size:

1. The number of instances (N) and the number of features per time point (m) have a significant impact on both model selection criteria and a not so strong influence in the difference in entropy and log-likelihood measures.
2. The time-series length (L) does not considerably affect none of the four measures.
3. As expected, with the increase of m, the number of instances (N) in a dataset also has to increase significantly for the method to select the true model, that is, to elicit the optimal early classification time point (n).
4. AIC is less dependent on N than MDL, but the latter identifies the true model more consistently than the first score.
5. The decision upon the early classification time point can be ambiguous; that is, the three main measures that compose the MCEC algorithm can propose distinct values of n.

References

1. Griffin, M.P.; Moorman, J.R. Toward the Early Diagnosis of Neonatal Sepsis and Sepsis-Like Illness Using Novel Heart Rate Analysis. *Pediatrics* **2001**, *107*, 97–104. [CrossRef] [PubMed]
2. Xing, Z.; Pei, J.; Yu, P.S. Early classification on time series. *Knowl. Inf. Syst.* **2012**, *31*, 105–127. [CrossRef]
3. Xing, Z.; Pei, J.; Yu, P.S.; Wang, K. Extracting Interpretable Features for Early Classification on Time Series. In Proceedings of the 2011 SIAM International Conference on Data Mining, Mesa, AZ, USA, 28–30 April 2011; pp. 247–258.
4. Ye, L.; Keogh, E.J. Time series shapelets: A new primitive for data mining. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 28 June–1 July 2009; pp. 947–956.
5. Mueen, A.; Keogh, E.J.; Young, N.E. Logical-shapelets: An expressive primitive for time series classification. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011; pp. 1154–1162.
6. Ghalwash, M.F.; Radosavljevic, V.; Obradovic, Z. Extraction of Interpretable Multivariate Patterns for Early Diagnostics. In Proceedings of the 2013 IEEE 13th International Conference on Data Mining, Dallas, TX, USA, 7–10 December 2013; pp. 201–210.
7. He, G.; Duan, Y.; Peng, R.; Jing, X.; Qian, T.; Wang, L. Early classification on multivariate time series. *Neurocomputing* **2015**, *149*, 777–787. [CrossRef]
8. Mori, U.; Mendiburu, A.; Keogh, E.J.; Lozano, J.A. Reliable early classification of time series based on discriminating the classes over time. *Data Min. Knowl. Discov.* **2017**, *31*, 233–263. [CrossRef]
9. Koller, D.; Friedman, N. *Probabilistic Graphical Models: Principles and Techniques*; MIT Press: Cambridge, MA, USA, 2009.
10. Carvalho, A.M.; Adão, P.; Mateus, P. Hybrid learning of Bayesian multinets for binary classification. *Pattern Recognit.* **2014**, *47*, 3438–3450. [CrossRef]
11. Carvalho, A.M.; Adão, P.; Mateus, P. Efficient Approximation of the Conditional Relative Entropy with Applications to Discriminative Learning of Bayesian Network Classifiers. *Entropy* **2013**, *15*, 2716–2735. [CrossRef]
12. Carvalho, A.M.; Roos, T.; Oliveira, A.L.; Myllymäki, P. Discriminative Learning of Bayesian Networks via Factorized Conditional Log-Likelihood. *J. Mach. Learning Res.* **2011**, *12*, 2181–2210.
13. Carvalho, A.M. Scoring Functions for Learning Bayesian Networks. INESC-ID Technical Report. 2009. Available online: http://www.lx.it.pt/~asmc/pub/talks/09-TA/ta_pres.pdf (accessed on 15 October 2019).
14. Barron, A.R.; Rissanen, J.; Yu, B. The Minimum Description Length Principle in Coding and Modeling. *IEEE Trans. Inf. Theory* **1998**, *44*, 2743–2760. [CrossRef]
15. Guerra, R.P.; Carvalho, A.M.; Mateus, P. Model selection for clustering of pharmacokinetic responses. *Comput. Methods Programs Biomed.* **2018**, *162*, 11–18. [CrossRef]
16. Sousa, M.; Carvalho, A.M. Polynomial-Time Algorithm for Learning Optimal BFS-Consistent Dynamic Bayesian Networks. *Entropy* **2018**, *20*, 274. [CrossRef]
17. Monteiro, J.L.; Vinga, S.; Carvalho, A.M. Polynomial-time algorithm for learning optimal tree-augmented dynamic Bayesian networks. In Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence, Amsterdam, The Netherlands, 12–16 July 2015; pp. 622–631.
18. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Auto. Control* **1974**, *19*, 716–723. [CrossRef]
19. Burnham, K.P.; Anderson, D.R. Multimodel inference: Understanding AIC and BIC in model selection. *Sociol. Methods Res.* **2004**, *33*, 261–304. [CrossRef]
20. Yang, Y. Can the Strengths of AIC and BIC Be Shared? *Biometrica* **2003**, *92*, 2003.
21. Vrieze, S.I. Model selection and psychological theory: A discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). *Psychol. Methods* **2012**, *17*, 228. [CrossRef] [PubMed]
22. Ridder, F.D.; Pintelon, R.; Schoukens, J.; Gillikin, D.P. Modified AIC and MDL model selection criteria for short data records. *IEEE Trans. Instrum. Meas.* **2005**, *54*, 144–150. [CrossRef]
23. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2006.

24. Amari, S. Information geometry on hierarchy of probability distributions. *IEEE Trans. Inf. Theory* **2001**, *47*, 1701–1711. [CrossRef]
25. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA Data Mining Software: An Update. *SIGKDD Explor. Newsl.* **2009**, *11*, 10–18. [CrossRef]
26. Bagnall, A.; Lines, J.; Bostrom, A.; Large, J.; Keogh, E. The Great Time Series Classification Bake Off: A Review and Experimental Evaluation of Recent Algorithmic Advances. *Data Min. Knowl. Discov.* **2016**. Avaiable online: http://timeseriesclassification.com/ (accessed on 15 October 2019). [CrossRef]
27. Mustafa., B. Multivariate Time Series Classification Data Sets. 2015. Available online: http://www.mustafabaydogan.com/files/viewcategory/20-data-sets.htmll (accessed on 15 October 2019).
28. Lin, J.; Keogh, E.J.; Wei, L.; Lonardi, S. Experiencing SAX: A novel symbolic representation of time series. *Data Min. Knowl. Discov.* **2007**, *15*, 107–144. [CrossRef]
29. Hu, B.; Rakthanmanon, T.; Hao, Y.; Evans, S.; Lonardi, S.; Keogh, E. Using the minimum description length to discover the intrinsic cardinality and dimensionality of time series. *Data Min. Knowl. Discov.* **2015**, *29*, 358–399. [CrossRef]
30. Olszewski, R. Generalized Feature Extraction for Structural Pattern Recognition in Time-Series Data. Ph.D Thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 2001.
31. Demsar, J. Statistical Comparisons of Classifiers over Multiple Data Sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30.
32. Buza, K.; Nanopoulos, A.; Schmidt-Thieme, L.; Koller, J. Fast Classification of Electrocardiograph Signals via Instance Selection. In Proceedings of the 2011 IEEE International Conference on Healthcare Informatics, Imaging and Systems Biology, HISB 2011, San Jose, CA, USA, 26–29 July 2011; pp. 9–16.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Journal of Cosmology
and Astroparticle
Physics

Entanglement-induced deviation from the geodesic motion in quantum gravity

To cite this article: Francisco Pipa *et al* JCAP09(2019)057

View the article online for updates and enhancements.

You may also like

- Strong gravity effects of rotating black holes: quasi-periodic oscillations
Alikram N Aliev, Göksel Daylan Esmer and Pamir Talazan
- Test particles in Kaluza–Klein models
João Baptista
- Infrared Horava–Lifshitz gravity coupled to Lorentz violating matter: a spectral action approach
D V Lopes, A Mamiya and A Pinzul

This content was downloaded from IP address 147.91.81.84 on 19/09/2025 at 14:21

Journal of Cosmology and Astroparticle Physics
An IOP and SISSA journal

Entanglement-induced deviation from the geodesic motion in quantum gravity

Francisco Pipa,[a,b] **Nikola Paunković**[c,d] **and Marko Vojinović**[e]

[a]Department of Philosophy, University of Kansas,
1445 Jayhawk Blvd., Wescoe Hall, Lawrence, KS 66045-7590, U.S.A.
[b]Departamento de Física, Instituto Superior Técnico, Universidade de Lisboa,
Avenida Rovisco Pais 1049-001, Lisboa, Portugal
[c]Instituto de Telecomunicações,
Av. Rovisco Pais 1, 1049-001, Lisbon, Portugal
[d]Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa,
Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
[e]Institute of Physics, University of Belgrade,
Pregrevica 118, 11080 Belgrade, Serbia

E-mail: franciscosapipa@gmail.com, npaunkov@math.tecnico.ulisboa.pt,
vmarko@ipb.ac.rs

Received August 6, 2019
Accepted August 29, 2019
Published September 26, 2019

Abstract. We study the derivation of the effective equation of motion for a pointlike particle in the framework of quantum gravity. Just like the geodesic motion of a classical particle is a consequence of classical field theory coupled to general relativity, we introduce the similar notion of an effective equation of motion, but starting from an abstract quantum gravity description. In the presence of entanglement between gravity and matter, quantum effects give rise to modifications of the geodesic trajectory, primarily as a consequence of the interference between various coherent states of the gravity-matter system. Finally, we discuss the status of the weak equivalence principle in quantum gravity and its possible violation due to the nongeodesic motion.

Keywords: gravity, quantum field theory on curved space, quantum gravity phenomenology

ArXiv ePrint: 1801.03207

© 2019 IOP Publishing Ltd and Sissa Medialab
https://doi.org/10.1088/1475-7516/2019/09/057

JCAP09(2019)057

Contents

1 Introduction **1**

2 Geodesic equation in general relativity **3**

3 Geodesic equation in quantum gravity **7**
- 3.1 Preliminaries and the setup 7
- 3.2 Effective covariant conservation equation 10
- 3.3 Effective equation of motion 12
- 3.4 Consistency of the approximation scheme 15

4 Status of the weak equivalence principle **16**
- 4.1 Definition and flavors of the equivalence principle 17
- 4.2 Equivalence principle and quantum theory 19
- 4.3 Universality, gravitational and inertial mass 23

5 Conclusions **25**
- 5.1 Summary of the results 25
- 5.2 Discussion of the results 26
- 5.3 Future lines of research 27

A Short review of the multipole formalism **28**

B Separable classical states **33**

C Phase of interference terms **33**

JCAP09(2019)057

1 Introduction

The formulation of the theory of quantum gravity (QG) is one of the most fundamental open problems in modern theoretical physics. In models of QG, as in any quantum theory, superpositions of states are allowed. In a tentative "theory of everything", which includes both gravity and matter at a fundamental quantum level, superpositions of product gravity-matter states are particularly interesting. Entangled states are highly nonclassical, and as such are especially relevant because they give rise to a drastically different behavior of matter from what one would expect based on classical intuition, as confirmed by numerous examples from the standard quantum mechanics (QM). Therefore, it is interesting to study such states in the context of a QG coupled to matter, in particular the Schrödinger cat-like states. Moreover, a recent study [1] suggests that physically allowed states of a gravity-matter system are generically entangled due to gauge invariance, providing additional motivation for our study.

In standard QM, entanglement is generically a consequence of the interaction. Nevertheless, there exist situations which give rise to entanglement even without interaction. For example, the Pauli exclusion principle in the case of identical particles generates entanglement without an interaction, giving rise to an effective force (also called the "exchange interaction"). We investigate in detail whether an entanglement between gravity and matter

could also be described as a certain type of an effective interaction, and if so, what are its aspects and details. In order to study this problem, we analyze the motion of a free test particle in a gravitational field. In general relativity (GR), this motion is described by a geodesic trajectory. However, we show that in the presence of the gravity-matter entanglement, the resulting effective interaction causes a deviation from a classical geodesic trajectory. In particular, we generalize the standard derivation of a geodesic equation from the case of classical gravity to the case of a full QG model, and derive the equation of motion for a particle which contains a non-geodesic term, reflecting the presence of the entanglement-induced effective interaction. The effects we discuss are purely quantum with respect to both gravity and matter, unlike previous studies of quantum matter in classical curved spacetime [2–5].

As a consequence of the modified equation of motion for a particle, we also discuss the status of the equivalence principle in the context of QG, and a possible violation of its weak flavor.

The paper is organized as follows. Section 2 is devoted to a review of the derivation of the geodesic equation in classical gravity, particularly in GR. The multipole formalism is employed and the geodesic equation for a particle is derived from the covariant conservation of the stress-energy tensor. In section 3 we generalize this procedure and derive our main results. Subsection 3.1 contains the general setup, the abstract quantum gravity framework that will be used, and the main assumptions. In subsection 3.2 we discuss the effective covariant conservation equation, which receives a correction to the classical one, due to the quantum gravity effects. In subsection 3.3 we put everything together and derive our main result — the effective equation of motion for a point particle, with the leading quantum correction. In subsection 3.4 we discuss the consistency of the assumptions that enter the approximation scheme used to derive the effective equation of motion. Section 4 is devoted to the discussion of the consequences of our results in the context of the weak equivalence principle. For the purpose of clarity, in subsection 4.1 we first provide the definitions of various flavors of the equivalence principle. Then, in subsection 4.2 we discuss the status of the equivalence principle in the context of quantum gravity and the results obtained in section 3. Subsection 4.3 provides further analysis of universality and equality between inertial and gravitational masses, in the context of the Newtonian approximation. Finally, section 5 contains our conclusions, discussion of the results and possible lines of further research. In the Appendix we give a short review of the multipole formalism used in the main text, with some mathematical details.

Our notation and conventions are as follows. We will work in the natural system of units in which $c = \hbar = 1$ and $G = l_p^2$, where l_p is the Planck length and G is the Newton's gravitational constant. By convention, the metric of spacetime will have the spacelike Lorentz signature $(-,+,+,+)$. The spacetime indices are denoted with lowercase Greek letters $\mu, \nu, \ldots$ and take the values $0, 1, 2, 3$. These can be split into the timelike index 0 and the spacelike indices denoted with lowercase Latin letters $i, j, k, \ldots$ which take the values $1, 2, 3$. The Lorentz-invariant metric tensor is denoted as $\eta_{\mu\nu}$. Quantum operators always carry a hat, $\hat{\phi}(x)$, $\hat{g}(x)$, etc. The parentheses around indices indicate symmetrization with respect to those indices, while brackets indicate antisymmetrization:

$$A_{(\mu\nu)} \equiv \frac{1}{2}\left(A_{\mu\nu} + A_{\nu\mu}\right) , \qquad A_{[\mu\nu]} \equiv \frac{1}{2}\left(A_{\mu\nu} - A_{\nu\mu}\right) .$$

Finally, we will systematically denote the values of functions with parentheses, $f(x)$, while functionals will be denoted with brackets, $F[\phi]$.

JCAP09(2019)057

2 Geodesic equation in general relativity

In the context of the classical theories of gravity, like GR, the question of deriving the geodesic equation for a particle has initially been studied by Einstein, Infeld and Hoffmann [6], Mathisson [7], Lubánski [8], Fock [9], and others. Slightly later, the question was revisited in the seminal paper by Papapetrou [10], with generalizations followed by a number of authors [11–22], developing the so-called *multipole formalism*, see the appendix A. Recently, the multipole formalism has been reformulated in a manifestly covariant language and extended from pointlike objects to strings, membranes and further to p-branes, with general equations of motion studied in Riemann and Riemann-Cartan spaces [23–28]. Today, the multipole formalism and the resulting classes of effective equations of motion have found applications in a wide range of topics, from string theory [29] to cosmology [30] to blackbrane dynamics [31–33] to elasticity and the studies of the shape of red blood cells in biological systems [34].

In this section we will demonstrate the application of the multipole formalism in its crudest *single pole* approximation, and employ it to derive the geodesic equation of motion for a point particle in classical Riemannian spacetime. The results presented in this section are well known in the literature, and illustrate the derivation procedure of the geodesic motion for a point particle. After reviewing the standard results in this section, in section 3 the same procedure will be utilized to study the quantum gravity case.

The derivation procedure is based on two main assumptions. The first assumption is that the matter fields have internal dynamics such that they form particle-like kink solutions which are stable (i.e., non-decaying) across the spacetime regions under consideration. If that is the case, one can employ the multipole formalism and expand the stress-energy tensor into a series of derivatives of the Dirac δ function as (see the appendix A for details):

$$T^{\mu\nu}(x) = \int_{\mathcal{C}} d\tau \left[B^{\mu\nu}(\tau) \frac{\delta^{(4)}(x - z(\tau))}{\sqrt{-g}} + \nabla_\rho \left(B^{\mu\nu\rho}(\tau) \frac{\delta^{(4)}(x - z(\tau))}{\sqrt{-g}} \right) + \dots \right] . \tag{2.1}$$

Here we assume that the stress-energy tensor of matter fields has nonzero value only near some timelike curve $\mathcal{C}$ represented by parametric equations $x^\mu = z^\mu(\tau)$, where τ is a parameter counting the points along the curve $\mathcal{C}$. In that case, the B-coefficients in the δ series will be smaller and smaller with each new term in the series. We introduce a series of smallness scales for the coefficients,

$$B^{\mu\nu} \sim \mathcal{O}_0 \,, \qquad B^{\mu\nu\rho} \sim \mathcal{O}_1 \,, \qquad B^{\mu\nu\rho\sigma} \sim \mathcal{O}_2 \,, \qquad \dots$$

such that one can consider the multipole scales to behave as

$$\mathcal{O}_0 \gg \mathcal{O}_1 \gg \mathcal{O}_2 \gg \dots \tag{2.2}$$

Next we choose to work in the so-called *single pole* approximation, in which all quantities of order $\mathcal{O}_1$ and higher can be neglected. It is also assumed that the typical radius of curvature of spacetime near the curve $\mathcal{C}$ will be large enough not to interfere in the internal dynamics of the matter fields along $\mathcal{C}$ and break the kink configuration apart. Physically speaking, the sequence of inequalities (2.2) states that one can systematically approximate the full solution of the matter field equations of motion by neglecting various degrees of freedom which describe the "size" and "shape" of the kink compared to its orbital motion (i.e., motion along the curve $\mathcal{C}$). Given this setup, in the single pole approximation the matter

JCAP09(2019)057

fields are in a configuration that looks like a point particle traveling along a worldline curve $\mathcal{C}$, and terms of order $\mathcal{O}_1$ and higher can be dropped from the stress-energy tensor, giving:

$$T^{\mu\nu}(x) = \int_{\mathcal{C}} d\tau\, B^{\mu\nu}(\tau) \frac{\delta^{(4)}(x - z(\tau))}{\sqrt{-g}}\,. \tag{2.3}$$

The second assumption is the validity of the local Poincaré invariance for the matter field equations. Namely, the classical action which describes the gravity-matter system can be generally written as

$$S[g, \phi] = S_G[g] + S_M[g, \phi]\,,$$

where g and ϕ denote gravitational and matter degrees of freedom, respectively, and it is generally considered to feature local Poincaré invariance. Our assumption is that the matter action S_M and the gravitational action S_G are invariant even taken separately. If this is the case, the Noether theorem gives us the covariant conservation of the stress-energy tensor of matter fields,

$$\nabla_\nu T^{\mu\nu} = 0\,. \tag{2.4}$$

Taken together, assumptions (2.3) and (2.4) are sufficient to establish two results:

(a) that the parametric functions $z(\tau)$ of the curve $\mathcal{C}$ satisfy the geodesic equation,

$$\frac{d^2 z^\lambda(\tau)}{d\tau^2} + \Gamma^\lambda{}_{\mu\nu} \frac{dz^\mu(\tau)}{d\tau} \frac{dz^\nu(\tau)}{d\tau} = 0\,, \tag{2.5}$$

where $\Gamma^\lambda{}_{\mu\nu}$ is the Christoffel connection for the background spacetime metric $g_{\mu\nu}$, and

(b) that the leading order coefficient $B^{\mu\nu}(\tau)$ in the stress-energy tensor for the particle has the form

$$B^{\mu\nu}(\tau) = m\, u^\mu(\tau) u^\nu(\tau)\,, \tag{2.6}$$

where $m \in \mathbb{R}\backslash\{0\}$ is an arbitrary constant parameter, while u^μ is the normalized tangent vector to the curve $\mathcal{C}$,

$$u^\mu \equiv \frac{dz^\mu(\tau)}{d\tau}\,, \qquad u^\mu u^\nu g_{\mu\nu} = -1\,.$$

In order to demonstrate these two statements, we start from (2.4), contract it with an arbitrary test function $f_\mu(x)$ of compact support, and integrate over the whole spacetime,

$$\int_{\mathcal{M}_4} d^4x \sqrt{-g}\, f_\mu \nabla_\nu T^{\mu\nu} = 0\,.$$

Then we perform the partial integration to move the covariant derivative from the stress-energy tensor to the test function. The boundary term vanishes since the test function has compact support, giving

$$\int_{\mathcal{M}_4} d^4x \sqrt{-g}\, T^{\mu\nu} \nabla_\nu f_\mu = 0\,.$$

Then we substitute (2.3), switch the order of integrations and perform the integral over spacetime $\mathcal{M}_4$, ending up with

$$\int_{\mathcal{C}} d\tau B^{\mu\nu} \nabla_\nu f_\mu = 0\,. \tag{2.7}$$

JCAP09(2019)057

The spacetime covariant derivative of the test function can be split into a component tangent to the curve $\mathcal{C}$ and a component orthogonal to it, in the following way. Using the identity

$$\delta^\lambda_\mu = -u^\lambda u_\mu + P^\lambda_{\perp\mu}\,, \tag{2.8}$$

where $-u^\lambda u_\mu$ and $P^\lambda_{\perp\mu}$ are projectors along u^μ and orthogonal to u^μ, respectively, we rewrite the derivative of f_ν as

$$\nabla_\nu f_\mu = -u_\nu \nabla f_\mu + f^\perp_{\nu\mu}\,, \tag{2.9}$$

where $\nabla \equiv u^\lambda \nabla_\lambda$ is the covariant derivative in the direction of the curve $\mathcal{C}$, while $f^\perp_{\nu\mu} \equiv P^\lambda_{\perp\nu}\nabla_\lambda f_\mu$ is a quantity orthogonal to the curve $\mathcal{C}$ with respect to its first index. Substituting (2.9) into (2.7), and performing another partial integration, we find

$$\int_{\mathcal{C}} d\tau \Big[f_\mu \nabla(B^{\mu\nu}u_\nu) + B^{\mu\nu} f^\perp_{\nu\mu} \Big] = 0\,,$$

where the boundary term again vanishes due to the compact support of the test function.

Given that the values of f_μ and $f^\perp_{\nu\mu}$ are both arbitrary and mutually independent along the curve $\mathcal{C}$, the coefficients multiplying them must each be zero. The first term gives us

$$\nabla(B^{\mu\nu}u_\nu) = 0\,, \tag{2.10}$$

while the second term, knowing that $f^\perp_{\nu\mu}$ is orthogonal to the curve $\mathcal{C}$ in its first index, gives

$$B^{\mu\nu}P^\lambda_{\perp\nu} = 0\,. \tag{2.11}$$

Focus first on (2.11). Knowing that $B^{\mu\nu}$ is symmetric, we can use (2.8) to decompose it into orthogonal and parallel components with respect to its two indices,

$$B^{\mu\nu} = B^{\mu\nu}_\perp + B^\mu_\perp u^\nu + B^\nu_\perp u^\mu + B u^\mu u^\nu\,,$$

where $B^{\mu\nu}_\perp$, $B^\mu_\perp$ and B are unknown coefficients, the first two being orthogonal to the curve $\mathcal{C}$ in all their indices. Substituting this expansion into (2.11), one finds that

$$B^{\mu\nu}_\perp = 0\,, \qquad B^\mu_\perp = 0\,,$$

leaving the scalar B as the only nonzero component of $B^{\mu\nu}$. Changing the notation from B to m, one obtains

$$B^{\mu\nu}(\tau) = m(\tau) u^\mu u^\nu\,. \tag{2.12}$$

This equation looks very similar to (2.6) but is still not equivalent to it, since the coefficient $m(\tau)$ is still not known to be a constant.

Next, focus on (2.10). Substituting (2.12), it reduces to

$$\nabla(m u^\mu) = 0\,. \tag{2.13}$$

Projecting onto the tangent direction u_μ and using the identity $u_\mu \nabla u^\mu = 0$, one obtains

$$\nabla m \equiv \frac{dm}{d\tau} = 0\,, \tag{2.14}$$

establishing that the parameter m is actually a constant. Given this, equation (2.13) reduces to

$$\nabla u^\mu = 0\,. \tag{2.15}$$

JCAP09(2019)057

Remembering that $\nabla \equiv u^\lambda \nabla_\lambda$ and expanding the covariant derivative, we see that this is the geodesic equation (2.5). Finally, (2.14) and (2.12) together give (2.6), which completes the proof of statements (a) and (b).

There are three general remarks one should make regarding the above procedure. The first remark is about the physical interpretation and properties of the free parameter m. Namely, it can be given the interpretation of the total mass of the particle — substituting (2.6) into the stress-energy tensor (2.3) and integrating the T^{00} component over the volume of the spatial hypersurface orthogonal to u^μ, one can easily verify that the total rest-energy of the matter fields at a given time is equal to m. Note, however, that the sign of m is not fixed to be positive. This is not surprising, since the covariant conservation equation (2.4) and the stress-energy tensor (2.3) do not contain any information (or assumption) about the positivity of energy. Instead, the positive energy condition $m > 0$ has to be established from the full matter field equations, which take into account the internal dynamics of the matter fields that make up the particle.

The second remark is about the metric $g_{\mu\nu}$ of the background geometry. When discussing the motion of a particle, the background geometry is usually assumed to be fixed, and backreaction of the gravitational field of the particle itself is not taken into account, leading to the notion of a "test particle". However, ignoring the backreaction is not a necessary assumption. Namely, one can take the full stress-energy tensor of the matter fields which form the kink solution (as opposed to the approximate single pole stress-energy tensor (2.3)), put it as a source into the Einstein's field equations and solve for the metric $g_{\mu\nu}$. The resulting metric does include the backreaction, and can then be reinserted into the geodesic equation for the particle. Note that this procedure is self-consistent, since the geodesic motion of the particle is a consequence of the covariant conservation equation (2.4) which is in turn itself a consequence of Einstein's field equations. Also note that the metric $g_{\mu\nu}$ obtained in this way does not necessarily give rise to the black hole geometry in the neighborhood of the particle. This is because the Schwarzschild radius of the kink may be (and usually is) much smaller than the scale $\mathcal{O}_1$ which defines the precision of the single pole approximation (2.3). A simple example would be the motion of a planet around the Sun — in the single pole approximation, the radius of the planet (itself far larger than the planet's gravitational radius) is considered to be of the order $\mathcal{O}_1$ and the planet is treated as a pointlike object, but the spacetime metric used in the geodesic equation can still take into account the planet's gravitational field in addition to the field of the Sun.

The third remark is about going beyond the single pole approximation. This has been studied in detail in the literature [10–21, 25–28], so here we merely point out the main physical interpretation. Namely, keeping the second term in the multipole expansion (2.1) physically amounts to giving the particle a nonzero "thickness", in the sense that its internal angular momentum can be considered nonzero. In the resulting equation of motion for the particle, this angular momentum couples to the spacetime curvature tensor, giving rise to a deviation from the geodesic motion. This can intuitively be understood as an effect of tidal forces acting across the scale of the kink's width, pushing it off the geodesic trajectory. Similarly, including quadrupole and higher order terms in (2.1) takes into account additional internal degrees of freedom of the kink, which also couple to spacetime geometry and produce a further deviation from geodesic motion.

The above review of the multipole formalism, and its application to the derivation of the geodesic equation in GR, will be used in the next section to discuss the corrections to the motion of a particle stemming from quantum gravity. As we shall see, these quantum

JCAP09(2019)057

corrections will give rise to additional terms in the effective equation of motion for a particle, pushing it slightly off the geodesic trajectory, even in the single pole approximation.

3 Geodesic equation in quantum gravity

In this section we discuss the motion of a particle within the framework of quantum gravity. The exposition is structured into four parts — first, we introduce the abstract quantum gravity formalism, and give some technical details about the description of the states. In the second part, we discuss the quantum version of the covariant conservation equation of the stress-energy tensor. In the third part we adapt the derivation presented in section 2 to the quantum formalism, and obtain the effective equation of motion for the particle. Finally, in the fourth part we discuss the self-consistency assumptions that go into the calculation.

3.1 Preliminaries and the setup

We work in the so-called generic abstract quantum gravity setup, as follows. Starting from the Heisenberg picture for the description of quantum systems, we assume that gravitational degrees of freedom are described by some gravitational field operators $\hat{g}(x)$, while matter degrees of freedom are described by matter field operators $\hat{\phi}(x)$, where x represents the coordinates of some point on a 4-dimensional spacetime manifold $\mathcal{M}_4$. Both sets of operators have their corresponding canonically conjugate momentum operators, $\hat{\pi}_g(x)$ and $\hat{\pi}_\phi(x)$, such that the usual canonical commutation relations hold. The total (kinematical) Hilbert space of the theory is $\mathcal{H}_{\text{kin}} = \mathcal{H}_G \otimes \mathcal{H}_M$, where the gravitational and matter Hilbert spaces $\mathcal{H}_G$ and $\mathcal{H}_M$ are spanned by the bases of eigenvectors for the operators $\hat{g}$ and $\hat{\phi}$, respectively. The total state of the system, $|\Psi\rangle \in \mathcal{H}_{\text{kin}}$, does not depend on x, in line with the Heisenberg picture framework.

There are several important points that need to be emphasized regarding the above setup. First, we do not explicitly state what are the fundamental degrees of freedom $\hat{g}$ for the gravitational field. They can be chosen in many different ways, giving rise to different models of quantum gravity. Since we aim to present the analysis of geodesic motion which is model-independent, we refrain from specifying what are the fundamental degrees of freedom $\hat{g}$. Instead, we merely assume that the operators describing the spacetime geometry, i.e., the metric, connection, curvature, etc., depend somehow on $\hat{g}$ and $\hat{\pi}_g$, and are expressible as operator functions in terms of them:

$$\hat{g}_{\mu\nu} = \hat{g}_{\mu\nu}(\hat{g}, \hat{\pi}_g)\,, \qquad \hat{\Gamma}^{\lambda}{}_{\mu\nu} = \hat{\Gamma}^{\lambda}{}_{\mu\nu}(\hat{g}, \hat{\pi}_g)\,, \qquad \hat{R}^{\lambda}_{\mu\nu\rho} = \hat{R}^{\lambda}_{\mu\nu\rho}(\hat{g}, \hat{\pi}_g)\,, \qquad \dots$$

When discussing these geometric operators, for simplicity we will usually not explicitly write their $(\hat{g}, \hat{\pi}_g)$-dependence.

Second, in order for any operator function to be well defined, some operator ordering has to be assumed. However, since we aim to work in an abstract model-independent QG formalism, we do not choose any particular ordering, but merely assume that one such ordering has been fixed. In a similar fashion, we also simply assume that all operators and spaces are well defined, convergent, and otherwise specified in enough mathematical detail to have a well defined and unique QG model. In a nutshell, our calculations are formal, in the sense that one should be able to repeat them in a detailed fashion if one is given a specific model of QG. This also means that our analysis and results should not depend on any of these details, but are rather common to a large class of QG models, and are based only on very few assumptions given above.

JCAP09(2019)057

Third, we employ a natural distinction between gravitational and matter degrees of freedom. Namely, whereas geometric operators such as metric, curvature, and so on, depend only on $\hat{g}$ and $\hat{\pi}_g$, matter operators like field strengths, stress-energy tensor, etc., will generically be operator functions of both $\hat{g}$, $\hat{\pi}_g$, and the fundamental matter degrees of freedom $\hat{\phi}$ and $\hat{\pi}_\phi$. In other words, we assume that the separation between gravity and matter present in the classical theory, described by an action of the form

$$S_{\text{total}}[g,\phi] = S_{\text{gravity}}[g] + S_{\text{matter}}[g,\phi]\,,$$

remains present also in the full quantum regime. That is to say, we assume that one can construct a theory of quantum gravity without matter fields, using only gravitational degrees of freedom g, so that this theory gives sourceless Einstein's equations of GR in the classical limit. Once such a pure-QG model has been constructed, we assume one can couple matter ϕ to it without changing the structure of the gravitational sector, obtaining the full QG model which features Einstein's equations with appropriate matter sources in the classical limit. While we do not consider this to be a strong assumption, we feel that it is nevertheless important to spell it out explicitly, since there may exist some QG models which fail to satisfy it, and our analysis may be inapplicable to such models.

After the introduction of the above conceptual setup, we turn to some more practical details. For the purpose of discussing geodesic motion, we are mostly interested in the effective classical theory of the abstract QG introduced above. To that end, the main objects of attention are *classical states* of gravity and matter, denoted by $|\Psi\rangle \in \mathcal{H}_G \otimes \mathcal{H}_M$. By classical, we mean that the "effective classical" values for the metric tensor and the matter stress-energy tensor, given by the expectation values of the corresponding operators

$$g_{\mu\nu} = \langle\Psi|\hat{g}_{\mu\nu}|\Psi\rangle\,, \qquad T_{\mu\nu} = \langle\Psi|\hat{T}_{\mu\nu}|\Psi\rangle \tag{3.1}$$

satisfy classical Einstein equations of the GR. A recent study suggests that physical states of gravity and matter are generically entangled [1]. For our analysis, we do not need to assume that the overall gravity-matter state is separable, and thus we will work with a generic state $|\Psi\rangle$ (see appendix B for the discussion of the separable case).

For the purpose of our paper, we will consider a *toy example state*, defined as

$$|\mathbf{\Psi}\rangle = \alpha|\Psi\rangle + \beta|\tilde{\Psi}\rangle\,, \tag{3.2}$$

where $|\tilde{\Psi}\rangle$ is some other classical state analogous to $|\Psi\rangle$, but giving different expectation values for the classical metric and stress-energy tensors:

$$\tilde{g}_{\mu\nu} = \langle\tilde{\Psi}|\hat{g}_{\mu\nu}|\tilde{\Psi}\rangle\,, \qquad \tilde{T}_{\mu\nu} = \langle\tilde{\Psi}|\hat{T}_{\mu\nu}|\tilde{\Psi}\rangle\,. \tag{3.3}$$

One can see that our toy-example state (3.2) is a Schrödinger-cat type of state, describing a coherent superposition of two classical configurations of gravitational and matter fields. It will become evident later on that qualitative conclusions of the paper do not depend on the fact that (3.2) has precisely two terms in the sum. Choosing the state with three, four or more terms will lead to analogous conclusions, although quantitative details of the computation may become technically more involved.

Given that (3.2) is a Schrödinger-cat type of state, there are some phenomenological restrictions on the values of the independent parameters β and $S \equiv \langle\Psi|\tilde{\Psi}\rangle$. Namely, in the ordinary experimental situations we basically never observe this kind of states, which

JCAP09(2019)057

means that the overall entangled state $|\mathbf{\Psi}\rangle$ looks pretty much like a classical state, say the state $|\Psi\rangle$. In other words, we want the fidelity between these two states to be large, $F(|\Psi\rangle, |\mathbf{\Psi}\rangle) = |\langle\Psi|\mathbf{\Psi}\rangle| \approx 1$. From (3.2) we obtain

$$\langle\Psi|\mathbf{\Psi}\rangle = \alpha + \beta S \equiv \kappa\,.$$

Define $|\Psi^\perp\rangle$, such that

$$|\tilde{\Psi}\rangle = S|\Psi\rangle + \epsilon|\Psi^\perp\rangle\,,$$

$\langle\Psi|\Psi^\perp\rangle = 0$ and $\epsilon = \sqrt{1-|S|^2}$. Thus,

$$F^2 = |\langle\Psi|\mathbf{\Psi}\rangle|^2 = 1 - \eta^2\,, \tag{3.4}$$

where we introduce the small parameter

$$\eta \equiv \beta\epsilon\,.$$

Here, we have used the normalization condition for the entangled state (3.2),

$$\langle\mathbf{\Psi}|\mathbf{\Psi}\rangle = \alpha^2 + \beta^2 + 2\alpha\beta\,\mathrm{Re}(S) = 1\,.$$

Given the above definitions for κ and η, we can rewrite the total state (3.2) as

$$|\mathbf{\Psi}\rangle = \kappa|\Psi\rangle + \eta|\Psi^\perp\rangle\,, \tag{3.5}$$

where $|\kappa|^2 = F^2 = 1-\eta^2$. The physical requirement of large fidelity implies that we study the limit $|\kappa| \approx 1$, $\eta \to 0$. We will therefore systematically expand the expectation values of all operators into power series in η, up to order $\mathcal{O}(\eta^2)$.

At this point we can evaluate the expectation values for the metric and stress-energy operators in the state (3.5), obtaining

$$\boldsymbol{g}_{\mu\nu} = \langle\mathbf{\Psi}|\hat{g}_{\mu\nu}|\mathbf{\Psi}\rangle = (1-\eta^2)g_{\mu\nu} + \eta^2\langle\Psi^\perp|\hat{g}_{\mu\nu}|\Psi^\perp\rangle + 2\eta\,\mathrm{Re}\left(\kappa\langle\Psi^\perp|\hat{g}_{\mu\nu}|\Psi\rangle\right)\,, \tag{3.6}$$

$$\boldsymbol{T}_{\mu\nu} = \langle\mathbf{\Psi}|\hat{T}_{\mu\nu}|\mathbf{\Psi}\rangle = (1-\eta^2)T_{\mu\nu} + \eta^2\langle\Psi^\perp|\hat{T}_{\mu\nu}|\Psi^\perp\rangle + 2\eta\,\mathrm{Re}\left(\kappa\langle\Psi^\perp|\hat{T}_{\mu\nu}|\Psi\rangle\right)\,. \tag{3.7}$$

It is easy to see that interference terms from the above expressions are generically nonvanishing. Indeed, even if, say, $\kappa\langle\Psi^\perp|\hat{g}_{\mu\nu}|\Psi\rangle$ were purely imaginary, a simple change of relative phase between $|\Psi\rangle$ and $|\tilde{\Psi}\rangle$ would give rise to a nontrivial real part. Namely, given a fixed choice of $|\tilde{\Psi}\rangle$, the set of choices for $|\Psi\rangle$ for which the interference term is purely imaginary is of measure zero compared to the full set of possible phase shifts of $|\Psi\rangle$. An analogous argument applies for $\kappa\langle\Psi^\perp|\hat{T}_{\mu\nu}|\Psi\rangle$ as well. For a detailed analysis, see appendix C.

Let us denote the metric and stress-energy interference terms as $\bar{g}_{\mu\nu}$ and $\bar{T}_{\mu\nu}$, respectively. Since we want to expand (3.6) and (3.7) into power series in η up to linear order, we can write

$$\bar{g}_{\mu\nu} \equiv 2\,\mathrm{Re}\left(\kappa\langle\Psi^\perp|\hat{g}_{\mu\nu}|\Psi\rangle\right) = h_{\mu\nu} + \mathcal{O}(\eta)\,, \tag{3.8}$$

$$\bar{T}_{\mu\nu} \equiv 2\,\mathrm{Re}\left(\kappa\langle\Psi^\perp|\hat{T}_{\mu\nu}|\Psi\rangle\right) = t_{\mu\nu} + \mathcal{O}(\eta)\,. \tag{3.9}$$

Here, $h_{\mu\nu}$ and $t_{\mu\nu}$ are η-independent parts of $\bar{g}_{\mu\nu}$ and $\bar{T}_{\mu\nu}$. Thus, we can finally write:

$$\boldsymbol{g}_{\mu\nu} = g_{\mu\nu} + \eta\,h_{\mu\nu} + \mathcal{O}(\eta^2)\,, \tag{3.10}$$

$$\boldsymbol{T}_{\mu\nu} = T_{\mu\nu} + \eta\,t_{\mu\nu} + \mathcal{O}(\eta^2)\,. \tag{3.11}$$

JCAP09(2019)057

In what follows, we will refer to the classical state $|\Psi\rangle$ as the *dominant state*, while the other classical state $|\tilde{\Psi}\rangle$ will be called the *sub-dominant state*. To justify this terminology, recall the above requirement (3.4) that the overall entangled state $|\boldsymbol{\Psi}\rangle$ looks like the classical state $|\Psi\rangle$, i.e., $F^2 = 1 - \eta^2$, with the parameter $\eta \equiv \beta\epsilon$ being small. Therefore, in the case $\beta \to 0$ and ϵ finite, the state $|\tilde{\Psi}\rangle$ enters (3.2) with a very small contribution, and is thus sub-dominant. On the other hand, in the case when β is finite and $\epsilon \to 0$, the states $|\Psi\rangle$ and $|\tilde{\Psi}\rangle$ are essentially indistinguishable, and their roles can be exchanged, as either can be considered sub-dominant to the other. By convention, we choose $|\tilde{\Psi}\rangle$ to again play the role of the sub-dominant state.

While in any quantum theory entangled states are allowed, note that when considering a product state of the gravity-matter system (i.e., the case $\eta = 0$), there is a danger that such a state may fail to be gauge invariant, as argued in [1]. So we need to introduce at least a small sub-dominant state, in order to ensure the gauge invariance of the total state. The simplest possible candidate state which describes the classical physics sufficiently well, and simultaneously stands a chance of being gauge invariant, is the genuinely entangled state (3.2), with $\beta \neq 0$ and $|\tilde{\Psi}\rangle \neq |\Psi\rangle$, leading to η being very small, but nonzero.

Regarding the effective entangled metric and stress-energy tensors (3.10) and (3.11), it is important to stress that they do not satisfy classical Einstein's equations of GR. Namely, we assume that Einstein's equations are separately satisfied by the metric and stress-energy tensors (3.1) coming from the classical state $|\Psi\rangle$, and by the metric and stress-energy tensors (3.3) coming from the other classical state $|\tilde{\Psi}\rangle$, as two different classical solutions:

$$R_{\mu\nu}(g) - \frac{1}{2} g_{\mu\nu} R(g) = 8\pi l_p^2 \, T_{\mu\nu} \,, \qquad R_{\mu\nu}(\tilde{g}) - \frac{1}{2} \tilde{g}_{\mu\nu} R(\tilde{g}) = 8\pi l_p^2 \, \tilde{T}_{\mu\nu} \,.$$

However, due to the nonlinearity of Einstein's equations, and due to the presence of the interference terms $h_{\mu\nu}$ and $t_{\mu\nu}$ in (3.10) and (3.11), quantities $\boldsymbol{g}_{\mu\nu}$ and $\boldsymbol{T}_{\mu\nu}$ do not satisfy Einstein's equations, as long as $\eta \neq 0$. This leads us to the following physical interpretation. First, it is natural to expand all quantities as corrections to the dominant classical configuration $(g_{\mu\nu}, T_{\mu\nu})$, including the equation of motion for a point particle. Second, as we shall see in the remainder of the text, given that $(\boldsymbol{g}_{\mu\nu}, \boldsymbol{T}_{\mu\nu})$ contains quantum gravity corrections through the interference terms, the presence of these quantum corrections in (3.10) and (3.11) will introduce an "effective force" term into the effective equation of motion for the particle. Finally, this effective force term will be pushing the particle off the geodesic trajectory defined by the classical dominant metric $g_{\mu\nu}$.

3.2 Effective covariant conservation equation

After the discussion of the general QG setup and the state (3.2), we move on to the discussion of the quantum analog of the covariant conservation equation (2.4). As in the classical theory, our basic assumption is that the matter sector of our QG model features local Poincaré invariance, i.e., that this symmetry is preserved at the quantum level. This assumption gives rise to a Gupta-Bleuler-like condition on the physical states, in the form

$$\langle \boldsymbol{\Psi} | \hat{\nabla}_\nu \hat{T}^{\mu\nu} | \boldsymbol{\Psi} \rangle = 0 \,, \tag{3.12}$$

where $\hat{\nabla}_\mu$ is the covariant derivative operator, defined by promoting the metric appearing in the Christoffel symbols into a corresponding operator. In general, the action of the stress-

JCAP09(2019)057

energy operator on the state $|\boldsymbol{\Psi}\rangle$ can be written[1] as

$$\hat{T}^{\mu\nu}|\boldsymbol{\Psi}\rangle = \boldsymbol{T}^{\mu\nu}|\boldsymbol{\Psi}\rangle + \boldsymbol{\Delta T}^{\mu\nu}|\boldsymbol{\Psi}^{\perp}\rangle\,, \tag{3.13}$$

where $\boldsymbol{T}^{\mu\nu}$ and $\boldsymbol{\Delta T}^{\mu\nu}$ are the expectation value and the uncertainty of the operator $\hat{T}^{\mu\nu}$ in the state $|\boldsymbol{\Psi}\rangle$, respectively,

$$\boldsymbol{T}^{\mu\nu} \equiv \langle\boldsymbol{\Psi}|\hat{T}^{\mu\nu}|\boldsymbol{\Psi}\rangle\,, \qquad \boldsymbol{\Delta T}^{\mu\nu} \equiv \sqrt{\langle\boldsymbol{\Psi}|\left(\hat{T}^{\mu\nu}\right)^2|\boldsymbol{\Psi}\rangle - \left(\langle\boldsymbol{\Psi}|\hat{T}^{\mu\nu}|\boldsymbol{\Psi}\rangle\right)^2}\,,$$

while $|\boldsymbol{\Psi}^{\perp}\rangle$ is some state orthogonal to $|\boldsymbol{\Psi}\rangle$. Note that the equation of the form (3.13) is completely general, holding for any stress-energy operator acting on an arbitrary state. Substituting (3.13) into (3.12), we obtain

$$\boldsymbol{\nabla}_\nu \boldsymbol{T}^{\mu\nu} + \langle\boldsymbol{\Psi}|\hat{\nabla}_\nu|\boldsymbol{\Psi}^{\perp}\rangle\boldsymbol{\Delta T}^{\mu\nu} = 0\,, \tag{3.14}$$

where $\boldsymbol{\nabla}_\nu$ is the expectation value of the operator $\hat{\nabla}_\nu$,

$$\boldsymbol{\nabla}_\nu \equiv \langle\boldsymbol{\Psi}|\hat{\nabla}_\nu|\boldsymbol{\Psi}\rangle\,.$$

At this point we need to make one more assumption. Namely, we assume that the error scale of the single pole approximation is bigger than the uncertainty of the stress-energy operator, $\boldsymbol{\Delta T}^{\mu\nu}$. Symbolically,

$$\mathcal{O}_1 \gtrsim \boldsymbol{\Delta T}^{\mu\nu}\,. \tag{3.15}$$

This means that in the single pole approximation we do not see the effects of the quantum fluctuations of matter fields. Intuitively, this is a reasonable assumption in most cases. For example, in the case of the kink solution describing the hydrogen atom, the scale on which one can detect quantum fluctuations (i.e., the Lamb shift effects) is much smaller than the size of the atom itself (i.e., the radius of the first Bohr orbit). Therefore, we expect that if our single pole approximation ignores the size of the atom itself, it also ignores the corresponding quantum fluctuations. An analogous assumption is made in relation to the uncertainty of the metric operator $\hat{g}_{\mu\nu}$,

$$\mathcal{O}_1 \gtrsim \boldsymbol{\Delta g}_{\mu\nu}\,, \tag{3.16}$$

given that the quantum gravity fluctuations can arguably also be ignored in the single pole approximation.

Applying (3.15) to (3.14), in the single pole approximation the second term can be dropped, leading to the effective classical covariant conservation equation,

$$\boldsymbol{\nabla}_\nu \boldsymbol{T}^{\mu\nu} = 0\,. \tag{3.17}$$

In a similar fashion, one can employ (3.16) to drop the off-diagonal components in the Christoffel symbol operators, leading to an effective classical expression

$$\boldsymbol{\Gamma}^{\lambda}{}_{\mu\nu} = \frac{1}{2}\boldsymbol{g}^{\lambda\sigma}\left(\partial_\mu \boldsymbol{g}_{\sigma\nu} + \partial_\nu \boldsymbol{g}_{\sigma\mu} - \partial_\sigma \boldsymbol{g}_{\mu\nu}\right)\,, \tag{3.18}$$

where $\boldsymbol{g}_{\mu\nu} \equiv \langle\boldsymbol{\Psi}|\hat{g}_{\mu\nu}|\boldsymbol{\Psi}\rangle$ is the effective classical metric and $\boldsymbol{g}^{\mu\nu}$ is its inverse matrix.

[1] Given any self-adjoint operator $\hat{A}$ and any state $|\Psi\rangle$, one can write

$$\hat{A}|\Psi\rangle = a|\Psi\rangle + b|\Psi^{\perp}\rangle\,,$$

where $\langle\Psi|\Psi^{\perp}\rangle \equiv 0$ and $a, b \in \mathbb{C}$. Multiplying this equation by $\langle\Psi|$ and by $\langle\Psi|\hat{A}$ from the left, one easily obtains that a and b are the expectation value and the uncertainty of the operator $\hat{A}$ in the state $|\Psi\rangle$, respectively.

JCAP09(2019)057

With effective classical expressions (3.17) and (3.18) in hand, we can now employ (3.10) and (3.11) to expand them into the dominant and correction parts. First we use (3.10) and $\boldsymbol{g}_{\mu\lambda}\boldsymbol{g}^{\lambda\nu} = \delta^{\nu}_{\mu}$ to find the inverse entangled metric $\boldsymbol{g}^{\mu\nu} = g^{\mu\nu} - \eta g^{\mu\rho}g^{\nu\sigma}h_{\rho\sigma} + \mathcal{O}(\eta^2)$, and then substitute into (3.18) to obtain

$$\boldsymbol{\Gamma}^{\lambda}{}_{\mu\nu} = \Gamma^{\lambda}{}_{\mu\nu} + \frac{\eta}{2}g^{\lambda\sigma}\left(\nabla_{\mu}h_{\sigma\nu} + \nabla_{\nu}h_{\sigma\mu} - \nabla_{\sigma}h_{\mu\nu}\right) + \mathcal{O}(\eta^2)\,, \tag{3.19}$$

where the Christoffel symbols in ordinary ∇_{μ} are defined with respect to the dominant classical metric $g_{\mu\nu}$. Then, expanding (3.17) into the form

$$\partial_{\nu}\boldsymbol{T}^{\mu\nu} + \boldsymbol{\Gamma}^{\mu}{}_{\sigma\nu}\boldsymbol{T}^{\sigma\nu} + \boldsymbol{\Gamma}^{\nu}{}_{\sigma\nu}\boldsymbol{T}^{\mu\sigma} = 0\,,$$

we substitute (3.11) and (3.19), and after a bit of algebra we rewrite it as:

$$\nabla_{\nu}T^{\mu\nu} + \eta\left[\nabla_{\nu}t^{\mu\nu} + T^{\sigma\nu}\left(\nabla_{\sigma}h^{\mu}{}_{\nu} - \frac{1}{2}\nabla^{\mu}h_{\nu\sigma}\right) + \frac{1}{2}T^{\mu\sigma}\nabla_{\sigma}h^{\nu}{}_{\nu}\right] + \mathcal{O}(\eta^2) = 0\,. \tag{3.20}$$

This equation is the one we sought out — it represents the analog of the classical covariant conservation equation (2.4), while taking into account the interference terms between the two classical states in (3.2), approximated to the linear order in η.

As a final step, (3.20) can be rewritten in a more compact form. For convenience, introduce the following shorthand notation (see our conventions from the last paragraph of the Introduction),

$$F^{\mu}{}_{\nu\sigma} \equiv \nabla_{(\sigma}h^{\mu}{}_{\nu)} - \frac{1}{2}\nabla^{\mu}h_{\nu\sigma}\,, \tag{3.21}$$

and also note that

$$F^{\nu}{}_{\nu\sigma} = \frac{1}{2}\nabla_{\sigma}h^{\nu}{}_{\nu} + \frac{1}{2}\nabla_{\nu}h^{\nu}{}_{\sigma} - \frac{1}{2}\nabla^{\nu}h_{\nu\sigma} = \frac{1}{2}\nabla_{\sigma}h^{\nu}{}_{\nu}\,,$$

so that, dropping the term $\mathcal{O}(\eta^2)$, equation (3.20) is rewritten as:

$$\nabla_{\nu}\left(T^{\mu\nu} + \eta t^{\mu\nu}\right) + 2\eta F^{(\mu}{}_{\nu\sigma}T^{\nu)\sigma} = 0\,. \tag{3.22}$$

The equation (3.22) represents the effective classical covariant conservation law for the stress-energy tensor, with the included quantum correction, represented to first order in η. It is the starting point for the remainder of our analysis, and replaces equation (2.4) in the derivation of the equation of motion for a point particle.

Finally, note that the quantum correction term in (3.22) has two distinct parts — one part comes from the quantum correction to the dominant classical stress-energy tensor, i.e., the interference term $t^{\mu\nu}$, while the other part comes from the quantum correction to the dominant classical metric, i.e., the interference term $h_{\mu\nu}$. This latter quantum correction enters through the Christoffel connection terms present in the covariant derivative. As we shall see in the next subsection, its presence will be crucial for the "force term" in the equation of motion for the particle, responsible for the deviation from the classical geodesic trajectory.

3.3 Effective equation of motion

We are now ready to derive the equation of motion for a particle in the single pole approximation, using the technique presented in section 2. However, instead of (2.4), we start from

JCAP09(2019)057

the effective covariant conservation law (3.22), which contains the quantum correction terms. Throughout, we assume the following relation of scales,

$$\mathcal{O}(\eta) > \mathcal{O}_1 \geq \mathcal{O}(\eta^2)\,.$$

In other words, we assume that the quantum correction terms linear in η are not smaller than the width of our particle, since otherwise one could simply ignore them and recover the classical geodesic motion for the particle.

Repeating the method of section 2, we begin by contracting (3.22) with an arbitrary test function $f_\mu(x)$ of compact support, and integrating over the whole spacetime,

$$\int_{\mathcal{M}_4} d^4x\sqrt{-g}\, f_\mu\Big[\nabla_\nu\left(T^{\mu\nu}+\eta t^{\mu\nu}\right)+2\eta F^{(\mu}{}_{\nu\sigma}T^{\nu)\sigma}\Big]=0\,.$$

We then perform the partial integration to move the covariant derivative from the stress-energy tensors to the test function. As before, the boundary term vanishes since the test function has compact support, giving

$$\int_{\mathcal{M}_4} d^4x\sqrt{-g}\,\Big[-\left(T^{\mu\nu}+\eta t^{\mu\nu}\right)\nabla_\nu f_\mu+2\eta F^{(\mu}{}_{\nu\sigma}T^{\nu)\sigma}f_\mu\Big]=0\,. \tag{3.23}$$

Now we need to model the dominant and correction parts of the stress-energy tensor. For the dominant part, it is straightforward to assume the single pole approximation, as was done in the classical case:

$$T^{\mu\nu}(x)=\int_{\mathcal{C}} d\tau B^{\mu\nu}(\tau)\frac{\delta^{(4)}(x-z(\tau))}{\sqrt{-g}}\,. \tag{3.24}$$

Regarding the correction term, we also use the single pole approximation,

$$t^{\mu\nu}(x)=\int_{\mathcal{C}} d\tau \bar{B}^{\mu\nu}(\tau)\frac{\delta^{(4)}(x-z(\tau))}{\sqrt{-g}}\,, \tag{3.25}$$

but one should note that in the case of $t^{\mu\nu}$ it is less obvious why this approximation is adequate, and requires some justification. However, in order to focus on the derivation of the particle equation of motion, for the moment we simply adopt (3.25), and postpone the analysis and the meaning of this approximation for subsection 3.4.

Then we substitute (3.24) and (3.25) into (3.23), switch the order of integrations and perform the integral over spacetime $\mathcal{M}_4$, ending up with

$$\int_{\mathcal{C}} d\tau\Big[-\left(B^{\mu\nu}+\eta\bar{B}^{\mu\nu}\right)\nabla_\nu f_\mu+2\eta F^{(\mu}{}_{\nu\sigma}B^{\nu)\sigma}f_\mu\Big]=0\,. \tag{3.26}$$

The next step is to employ the identity (2.9) to separate the tangential and orthogonal components of the derivative of the test function. Substituting it into (3.26), and performing another partial integration, we find

$$\int_{\mathcal{C}} d\tau\Big[\left(B^{\mu\nu}+\eta\bar{B}^{\mu\nu}\right)f^{\perp}_{\nu\mu}+\Big[\nabla\left(B^{\mu\nu}u_\nu+\eta\bar{B}^{\mu\nu}u_\nu\right)-2\eta F^{(\mu}{}_{\nu\sigma}B^{\nu)\sigma}\Big]f_\mu\Big]=0\,,$$

where the boundary term again vanishes due to the compact support of the test function.

JCAP09(2019)057

After these transformations, we make use of the same argument that both f_μ and $f^\perp_{\nu\mu}$ are arbitrary and mutually independent along the curve $\mathcal{C}$, concluding that the coefficients multiplying them must each be zero. The first term gives us

$$\nabla\left(B^{\mu\nu}u_\nu + \eta\bar{B}^{\mu\nu}u_\nu\right) - 2\eta F^{(\mu}{}_{\nu\sigma}B^{\nu)\sigma} = 0\,, \tag{3.27}$$

while the second term, knowing that $f^\perp_{\nu\mu}$ is orthogonal to the curve $\mathcal{C}$ in its first index, gives

$$\left(B^{\mu\nu} + \eta\bar{B}^{\mu\nu}\right)P^\lambda_{\perp\nu} = 0\,.$$

As in the previous case, given that $B^{\mu\nu}$ and $\bar{B}^{\mu\nu}$ are symmetric, one can decompose them into tangential and orthogonal components using (2.8), and then from (2.11) read off that all orthogonal components must be zero, concluding that

$$B^{\mu\nu} + \eta\bar{B}^{\mu\nu} = (B + \eta\bar{B})u^\mu u^\nu \equiv m(\tau)u^\mu u^\nu\,, \tag{3.28}$$

where again we emphasized that the parameter m may depend on the particle's proper time τ.

Next, substituting this into (3.27) and neglecting the term $\mathcal{O}(\eta^2)$, we obtain

$$\nabla\left(mu^\mu\right) + \eta m u^\sigma\left(F^\mu{}_{\nu\sigma}u^\nu + F^\nu{}_{\nu\sigma}u^\mu\right) = 0\,. \tag{3.29}$$

Projecting onto the tangent direction u_μ and using the identity $u_\mu\nabla u^\mu = 0$, one obtains

$$\nabla m = \eta m u^\sigma\left(u^\nu u_\lambda F^\lambda{}_{\nu\sigma} - F^\nu{}_{\nu\sigma}\right)\,, \tag{3.30}$$

establishing that, in contrast to the classical case, here the parameter m fails to be constant. Substituting this back into the equation (3.29), after some simple algebra we obtain

$$\nabla u^\mu + \eta u^\nu u^\sigma P^\mu_{\perp\lambda}F^\lambda{}_{\nu\sigma} = 0\,,$$

where the parameter m cancels out of the equation. As a final step, introducing the shorthand notation $F^\mu_{\perp\nu\sigma} \equiv P^\mu_{\perp\lambda}F^\lambda{}_{\nu\sigma}$, we can rewrite the equation of motion in its final form

$$\nabla u^\mu + \eta u^\nu u^\sigma F^\mu_{\perp\nu\sigma} = 0\,. \tag{3.31}$$

The presence of the orthogonal projector in the second term should not be surprising. Namely, since acceleration must always be orthogonal to the velocity, the second term in the equation must also be orthogonal to velocity, and this is guaranteed by the presence of the orthogonal projector.

Equations (3.28), (3.30) and (3.31) are the main result of this paper, and we discuss them in turn. Equation (3.28) determines the structure of the stress-energy tensor describing the point particle, as a function of tangent vectors of its world line and a scalar parameter $m(\tau)$. Formally, it has the same form as its classical counterpart (2.12), and provisionally the parameter m may be even called *effective mass*. Namely, in the rest frame of the particle, integration of the $\boldsymbol{T}^{00}$ component of the entangled stress-energy tensor over the 3-dimensional spatial hypersurface can be interpreted as the total rest-energy of the kink configuration of fields that represents the particle. This terminology is of course provisional, since all these notions are merely a part of the semiclassical approximation of the full quantum gravity description.

JCAP09(2019)057

Equation (3.30) determines the proper time evolution of the parameter $m(\tau)$. In contrast to the classical case, where $m(\tau)$ was determined by (2.14) to be a constant, here we see that its time derivative is proportional to (covariant derivatives of) the interference term $h_{\mu\nu}$ between the dominant and sub-dominant classical geometry, via (3.21). If one puts $\eta = 0$, (3.30) reduces to the classical case, as expected. The interference between the two geometries gives rise to an effective force that is responsible for the change in time of the particle's effective mass. Since the particle is (effectively) not isolated, its total energy is therefore not conserved, in the sense of equation (3.30).

Finally, and most importantly, equation (3.31) represents the effective equation of motion of the particle, determining its world line. It has the form of the classical geodesic equation (2.15) with an additional correction term proportional to η and to the interference term $h_{\mu\nu}$. This additional term represents an *effective force*, pushing the particle off the classical geodesic trajectory. It is analogous to the notion of the "exchange interaction" force in molecular physics, in the region where the wavefunctions of the two electrons in a molecule overlap.

In our case, however, the force term is determined by the interference between the two classical spacetime and matter configurations superposed in the state (3.2), and in particular by the off-diagonal components of the metric operator $\hat{g}_{\mu\nu}$, see (3.8). It is thus a *pure quantum gravity effect*, a consequence of the nontrivial structure of the metric operator. Of course, the detailed properties and the magnitude of the force term depend on the choice of the two classical gravity-matter configurations and on the details of the quantization of the gravitational field.

3.4 Consistency of the approximation scheme

Regarding the analysis and the derivation of the effective equation of motion for a particle discussed in the previous subsection, there is one issue that we should reflect on. It is related to the additional consistency conditions that stem from our assumption that the quantum correction to the stress-energy tensor is approximated with a single pole term (3.25).

Namely, the two stress-energy tensors that enter the derivation of the effective equation of motion — the classical dominant stress-energy tensor $T^{\mu\nu}$, and the interference stress-energy tensor $t^{\mu\nu}$ — can in general be written in the single pole approximation as:

$$T^{\mu\nu} = \int_{\mathcal{C}} d\tau\, B^{\mu\nu}(\tau) \frac{\delta^{(4)}(x - z(\tau))}{\sqrt{-g}} + \mathcal{O}_1(T)\,, \tag{3.32}$$

$$t^{\mu\nu} = \int_{\mathcal{C}} d\tau\, \bar{B}^{\mu\nu}(\tau) \frac{\delta^{(4)}(x - z(\tau))}{\sqrt{-g}} + \mathcal{O}_1(t)\,. \tag{3.33}$$

Note that we have introduced two different $\mathcal{O}_1$ scales, one for each tensor. This is because, although we assume that both can be expanded into the δ series around the same curve $\mathcal{C}$, each tensor may have different "width", or in other words, the two configurations of matter fields may be such that they can be well approximated with a single pole term up to a priori two different $\mathcal{O}_1$ scales. In particular, if one chooses the $\mathcal{O}_1$ scale to write $T^{\mu\nu}$ in a single pole approximation, $\mathcal{O}_1 = \mathcal{O}_1(T)$, it is not obvious that $t^{\mu\nu}$ can also be approximated by a single pole term, compared to the same scale, and vice versa. Therefore, there is an assumption about the relationship between scales that we have made when we used expressions (3.24) and (3.25) in the derivation of the effective equation of motion.

JCAP09(2019)057

Looking at the structure of the equation (3.23) into which (3.24) and (3.25) have been substituted, the consistency condition for the approximation scheme can be written as

$$\mathcal{O}_1 \equiv \mathcal{O}_1(T) \geq \eta \mathcal{O}_1(t)\,. \tag{3.34}$$

In particular, if this inequality were not valid, the dipole term in (3.33) would contribute to (3.23) with a magnitude comparable to the single pole term of (3.32), and it would be inconsistent to ignore it in the derivation of the effective equation of motion.

The consistency condition (3.34) can be rewritten into a more explicit form. Substituting (3.33) and (3.9) into (3.34), we get

$$\mathcal{O}_1(T) \geq \eta \left[2\,\mathrm{Re}\left(\langle \Psi^\perp | \hat{T}_{\mu\nu} | \Psi \rangle \right) - \int_{\mathcal{C}} d\tau\, \bar{B}^{\mu\nu}(\tau) \frac{\delta^{(4)}(x - z(\tau))}{\sqrt{-g}} \right] .$$

In addition, one can use (3.28) and (3.32) to eliminate the coefficient $\bar{B}^{\mu\nu}$ in favor of $T^{\mu\nu}$ and $m(\tau)$, which are arguably more observable, obtaining

$$\mathcal{O}_1(T) \geq 2\eta\,\mathrm{Re}\left(\langle \Psi^\perp | \hat{T}_{\mu\nu} | \Psi \rangle \right) + T^{\mu\nu} - \int_{\mathcal{C}} d\tau\, m(\tau) u^\mu u^\nu \frac{\delta^{(4)}(x - z(\tau))}{\sqrt{-g}}\,. \tag{3.35}$$

This inequality should be interpreted as follows. Given an explicit model of quantum gravity, and within it an explicit configuration of matter fields that make up a particle, one can estimate all three quantities on the right-hand side of (3.35), namely the off-diagonal components of the stress-energy operator, its expectation value in the dominant classical state, and the total mass of the particle, respectively. Then, the consistency condition (3.35) gives a lower bound on the scale $\mathcal{O}_1$, which represents an estimate of the error when discussing the effective equation of motion for the particle. In other words, the equation of motion can be considered to be approximately valid only across scales much larger than the $\mathcal{O}_1$ scale, bounded from below by inequality (3.35).

Finally, if one needs better precision than the scale determined by (3.35), one should take into account the dipole term in (3.33) and rederive a more precise form of the equation of motion. Still better precision would be obtained by including the dipole term in (3.32), which would amount to the equation of motion in the full pole-dipole approximation, and so on.

4 Status of the weak equivalence principle

In light of the results of section 3, it is important to discuss the status of the equivalence principle (EP). Throughout the literature, one can find various different formulations of EP, in various flavors such as weak, medium-strong, strong, and so on (see [35, 36] for a review, and [2–5, 37, 38] for various examples). Often these formulations and flavors are interpretation-dependent, and it is not always clear whether they are mutually equivalent or not, and what are the underlying assumptions and definitions used to express them.

Needless to say, such situation is less than satisfactory [35, 36], and in order to circumvent it, in this section we opt to specify one particular definition of the weak and strong equivalence principles (WEP and SEP, respectively) and to use this definition in the rest of the text. We do not aspire to claim that our definition is either equivalent to, or in any sense better than, other definitions present in the literature, and may not even correspond to the usual usage of the terminology. But for the purpose of clarity, it is prudent to fix one definition and stick to it. Therefore, in light of the results obtained in section 3, in this section we discuss the status of WEP defined as below.

JCAP09(2019)057

4.1 Definition and flavors of the equivalence principle

The purpose of the equivalence principle is to prescribe the coupling of matter to gravity [39]. Its precise formulation therefore depends on the particular choice of the gravitational and matter degrees of freedom which one uses to describe matter and gravity. For the purpose of this paper, we assume that the classical limit of quantum gravity corresponds to general relativity, which means that in this limit the fundamental gravitational degrees of freedom give rise to a nonflat spacetime metric. Given any choice of the gravitational degrees of freedom that belong to this class, in the classical framework one can formulate the equivalence principle as a two-step recipe to couple matter to gravity (we will discuss the quantum framework in subsection 4.2).

Start from the classical equation of motion for matter degrees of freedom in flat spacetime, written symbolically as

$$\mathcal{D}_{\text{flat}}[\phi, \eta_{\mu\nu}] = 0\,, \tag{4.1}$$

where ϕ denotes the matter degrees of freedom, $\eta_{\mu\nu}$ is the Minkowski metric, while $\mathcal{D}_{\text{flat}}$ is an appropriate functional describing the equation of motion for ϕ in flat spacetime and is assumed to be local. Given this equation of motion, couple it to gravity as follows:

1. Rewrite the equation of motion in a manifestly diffeomorphism-invariant form, typically by a change of variables to a generic curvilinear coordinate system,

   $$\mathcal{D}_{\text{curvilinear}}[\phi, g^{(0)}_{\mu\nu}] = 0\,,$$

   where $g^{(0)}_{\mu\nu}$ is still the flat spacetime metric, appropriately transformed from $\eta_{\mu\nu}$, and similarly for $\mathcal{D}_{\text{curvilinear}}$.

2. Promote the curvilinear equation of motion to the equation of motion in curved spacetime by replacing the flat spacetime metric $g^{(0)}_{\mu\nu}$ with an arbitrary metric $g_{\mu\nu}$,

   $$\mathcal{D}_{\text{curvilinear}}[\phi, g_{\mu\nu}] = 0\,,$$

   thereby specifying the equation of motion for matter coupled to gravity.

The first step describes the matter equation of motion from a perspective of a generic curvilinear (or "arbitrarily accelerated") coordinate system, reflecting the principle of *general relativity*. The second step simply promotes that same equation to curved spacetime as it stands, with no additional coupling of any kind. This can be loosely formulated as a statement of *local equivalence between gravity and acceleration*, which is how the EP historically got its name. Also, note that these two steps operationally correspond to the standard *minimal coupling* prescription [39].

It is important to stress the *local* nature of EP, which manifests itself in the assumption that the initial equation of motion (4.1) is local, and that the EP essentially does not change it at all, at any given point in spacetime. This has one important implication — the gravitational degrees of freedom manifest themselves only through *nonlocal measurements*, as tidal effects induced by spacetime curvature. We will return to this point and comment more on it later in the text.

Depending on the further specification of the matter degrees of freedom, one can distinguish between various flavors of the EP. For example, if one talks about the mechanics of point particles, one can start from the Newton's first law of motion, which states that

JCAP09(2019)057

in the absence of any forces, a particle has a straight-line trajectory in Minkowski spacetime. According to the step 1 above, the differential equation for a straight line in a generic curvilinear coordinate system is the geodesic equation,

$$\frac{d^2 z^\lambda}{d\tau^2} + \Gamma^\lambda_{(0)\mu\nu} \frac{dz^\mu}{d\tau} \frac{dz^\mu}{d\tau} = 0\,,$$

where the index (0) on the Christoffel symbol indicates that it is calculated using the metric $g^{(0)}_{\mu\nu}$, which is obtained by a curvilinear coordinate transformation from the Minkowski metric $\eta_{\mu\nu}$. Then, according to step 2, one again writes the same equation, only dropping the requirement of flat spacetime metric,

$$\frac{d^2 z^\lambda}{d\tau^2} + \Gamma^\lambda{}_{\mu\nu} \frac{dz^\mu}{d\tau} \frac{dz^\mu}{d\tau} = 0\,,$$

so that this time the Christoffel symbol is calculated using an arbitrary metric $g_{\mu\nu}$, and now encodes the interaction with the gravitational degrees of freedom. So one starts from the Newton's first law of motion for a particle in the absence of the gravitational field, and ends up with a geodesic equation in the presence of the gravitational field. We define this flavor of the EP as the *weak equivalence principle* (WEP).

Instead of mechanical particles, one can study matter degrees of freedom described by a field theory. For example, if one starts from the equation of motion for a single real scalar field,

$$(\eta^{\mu\nu}\partial_\mu\partial_\nu - m)\,\phi = 0\,,$$

according to the step 1 of the EP, one can rewrite it in a general curvilinear coordinate system as

$$\left(g^{\mu\nu}_{(0)}\nabla_\mu\nabla_\nu - m\right)\phi = 0\,,$$

where the Christoffel symbol inside the covariant derivative is again calculated using the flat-space metric $g^{(0)}_{\mu\nu}$. Then, according to step 2 of the EP, this equation is promoted to curved spacetime as it stands, leading to

$$(g^{\mu\nu}\nabla_\mu\nabla_\nu - m)\,\phi = 0\,,$$

where now the covariant derivative is given with respect to an arbitrary metric $g_{\mu\nu}$ describing curved spacetime. Thus one arrives to the equation of motion for a scalar field coupled to gravity. We define this flavor of the EP as the *strong equivalence principle* (SEP).

So in short, WEP is a statement about mechanical systems such as particles and small bodies, while SEP is a statement about fields. We emphasize again that the above definitions may or may not correspond to what is known in other literature as WEP and SEP, depending on the particular source one compares our definitions to. For example, one can often find a definition of WEP as a statement about equality of inertial and gravitational masses. As another example, one can also find a definition of WEP as Galileo's statement that the acceleration of a particle due to the gravitational field is independent of the particle's internal details such as mass or chemical composition, a property also called *universality*, emphasizing the fact that gravitation interacts with all types of particles in the same way. For an excellent review of the various formulations and flavors of EP present in the literature, see [35].

In relation to these alternative formulations of WEP, one should note two comments. First, while the notion of "gravitational mass" may be useful in the context of Newtonian

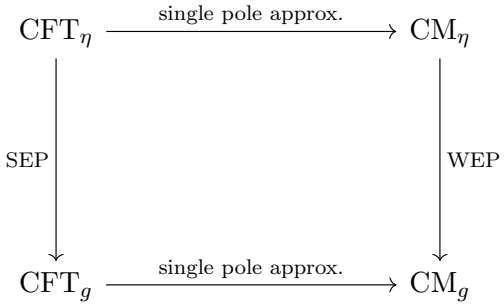
JCAP09(2019)057

theory, in frameworks such as GR it is not useful, since the source in Einstein's equations is the whole stress-energy tensor, rather than any particular mass-like parameter. This renders any definition of WEP which relies on the notion of the gravitational mass unsuitable for analysis in a fundamental QG framework. Second, one can argue (see for example [35]) that the property of universality is implicitly present even without gravitational interaction, in the Newton's first law of motion. Namely, the first Newton's law can be formulated more precisely as follows: in the absence of any forces, a particle has a straight-line trajectory in Minkowski spacetime, *regardless of its internal details such as mass or chemical composition.* The Newton's first law is never spelled out in this way in textbooks, making room for a point of view that universality has something to do with gravity or the EP. However, if one accepts our definition of WEP given above, it is more natural to say that universality is a property of Newtonian mechanics, and is merely *being preserved* by the WEP when one lifts the straight-line equation of motion to curved spacetime. So from this point of view, one should arguably say that WEP is merely *compatible* with universality, rather than equivalent to it.

Given all these reasons, and despite the fact that these alternative definitions of WEP may be suitable in various other contexts, they are not quite adequate for the analysis given in this paper. We therefore choose to retain our own definition of WEP, while the principles of universality and equality between inertial and gravitational mass will be called as such. They are discussed in more detail in subsection 4.3.

4.2 Equivalence principle and quantum theory

Adopting the above definitions of WEP and SEP, it is important to discuss their relationship. From the perspective of the classical field theory (CFT), the notion of a particle can be introduced as a localized kink-like configuration of matter fields, described as a solution of the (usually quite complicated) matter field equations. One can then employ the apparatus of multipole formalism and describe the evolution of this kink configuration in the single pole approximation, as was discussed in section 2. Using this method, one can recover the equation of motion for a particle in classical mechanics (CM) as an approximation of the field theory. Moreover, all this can be done before or after the application of the EP, leading to the following diagram:



Here the indices η and g indicate that equations of motion in a given theory are written in flat and in curved spacetime, respectively.

The question whether this diagram commutes is nontrivial. Namely, on one hand, one can start from a flat-space classical field theory, approximate it to derive the equations of motion for a particle in flat-space classical mechanics, and then invoke WEP to reach classical mechanics coupled to gravity. On the other hand, one can first invoke SEP to couple matter to gravity at the field theory level, and then approximate it to derive the equation of motion

JCAP09(2019)057

for a particle in curved spacetime. A priori, there is no guarantee that one will reach the same equation of motion for a particle in curved spacetime using both methods.

It is in fact the existence of the local Poincaré symmetry that leads to the commutativity of the diagram. Namely, as was discussed in section 2, in the curved spacetime local Poincaré symmetry gives rise to the covariant conservation equation for the stress-energy tensor of matter fields, and this is all one needs to reach the geodesic equation as an equation of motion for the particle, in the sense that one does not need to know the full matter field equations in curved spacetime. This establishes the $\langle \text{SEP} \to \text{single pole} \rangle$ path of the diagram. On the other hand, in flat spacetime one can also perform the calculation of section 2, this time using the ordinary (noncovariant) conservation equation for the stress-energy tensor, which is a consequence not of the local, but rather of the global Poincaré invariance of Minkowski spacetime. Repeating the calculation of section 2 with the symbolic substitutions $g \to \eta$ and $\nabla \to \partial$, it is not hard to conclude that one will obtain the equation of motion for a straight line in flat spacetime, again without knowing all details of the full matter field equations in flat spacetime. Then, applying WEP as discussed in subsection 4.1, one reaches the geodesic equation in curved spacetime. This establishes the $\langle \text{single pole} \to \text{WEP} \rangle$ path of the diagram, concluding that the resulting equation of motion for the particle is the same in both cases, i.e., that the diagram commutes.

Let us also note that, going beyond the single pole approximation, WEP is known to be violated, with SEP remaining valid. For example, in the pole-dipole approximation, it is well known that the analogous diagram

$$\begin{array}{ccc} \text{CFT}_\eta & \xrightarrow{\text{pole-dipole approx.}} & \text{CM}_\eta \\ \Big\downarrow \text{SEP} & & \Big\downarrow \text{WEP} \\ \text{CFT}_g & \xrightarrow{\text{pole-dipole approx.}} & \text{CM}_g \end{array}$$

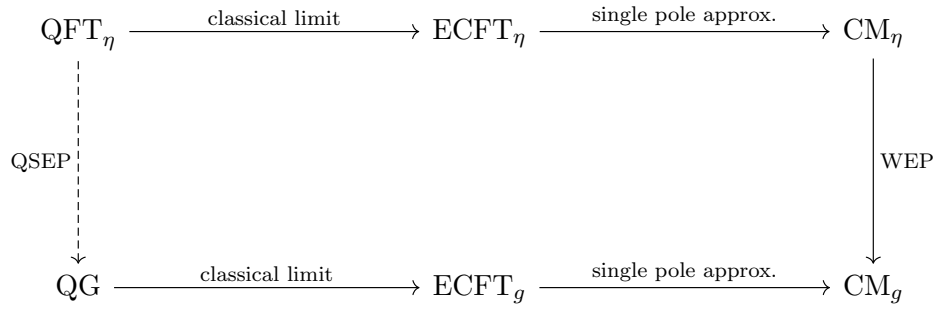
fails to commute. Namely, the $\langle \text{SEP} \to \text{pole-dipole} \rangle$ path leads to an effective equation of motion for the particle in which there is an explicit coupling of the particle's total angular momentum to the spacetime curvature [10]. On the other hand, the $\langle \text{pole-dipole} \to \text{WEP} \rangle$ path produces the equation of motion without the curvature term. Thus, in the pole-dipole approximation, WEP fails to reproduce the correct equation of motion, since the particle is coupled to gravity in a nonminimal way, in spite of the fact that the fields which make up the particle are still minimally coupled to gravity, in line with SEP. Of course, this situation is to be expected, given that in the pole-dipole approximation the particle is no longer completely pointlike, and the coupling of the angular momentum to the curvature can be understood as a tidal effect of gravity across the "width" of the particle. On the other hand, one can instead argue that it would be wrong to apply WEP to the pole-dipole equation of motion for a particle. Namely, despite the fact that the latter is formally still local, it describes an object that is "less-than-perfectly pointlike", in the sense that its stress-energy tensor is proportional not only to a δ function but also to its derivative. From that point of view, one should not be allowed to apply the two-step prescription of EP defined above. Either way, the bottom line is that one can either declare WEP as violated or as inapplicable beyond the single pole

JCAP09(2019)057

approximation, but it cannot be declared as valid. This results in the noncommutativity of the above diagram.

Let us now turn to the quantum theory. Starting first from some quantum field theory (QFT_η) which describes the fundamental matter fields in Minkowski spacetime, one can take its classical limit, giving rise to some effective classical field theory (ECFT_η). Then, assuming that the latter features kink solutions, one can describe those using the single pole approximation, leading to classical mechanics (CM_η) of the corresponding particles. Finally, applying WEP one couples those particles to gravity. The resulting equation of motion will always be a geodesic equation, assuming that the initial QFT and all subsequent approximations respect the global Poincaré invariance of Minkowski spacetime. This symmetry guarantees the conservation of the stress-energy tensor of the matter fields throughout the sequence of approximations, leading invariably to the geodesic equation of motion for the particle.

On the other hand, it is arguably more appropriate to take an alternative, more fundamental route — start from some fundamental quantum gravity (QG) model, and take the classical limit leading to some effective classical field theory (ECFT_g) for both matter and gravity. Then, again assuming that this theory features kink solutions, employ the single pole approximation to obtain the classical mechanics for the particle in the gravitational field (CM_g). Note that this is in fact precisely the program that was performed in section 3, leading to the non-geodesic equation of motion (3.31) for the particle. In effect, one can conclude that the following diagram fails to commute:



As a side comment, note that the dashed QSEP arrow represents some hypothetical map leading from a QFT in Minkowski spacetime to a full-blown model of QG, according to a notion that might be called a "quantum strong equivalence principle". It is unclear whether such a principle exists or not, let alone what its formulation is supposed to be, even if one is given precisely defined models of QFT and QG in question. We introduce it here simply for completeness, speculating that such a notion should exist, as a generalization of SEP from classical to quantum physics. It is also convenient to introduce it, in order to close the diagram and discuss its commutativity.

It is important to stress the reason why this diagram does not commute. Recalling the details of section 3, the local Poincaré symmetry is assumed to be respected at the fundamental level of QG and onwards, just like in the classical case. Moreover, the single pole approximation is used, avoiding any nonminimal coupling of the tidal forces that may be present. And yet, in spite of all that, the resulting equation of motion is not a geodesic. Looking at the equation of motion (3.31), the reason for this is the nontrivial interference between classical states describing two classical configurations of matter, and more importantly, of gravity. In other words, the deviation from the geodesic motion is a *pure quantum gravity effect* — it is not present in the classical case, nor in the case of quantum matter in

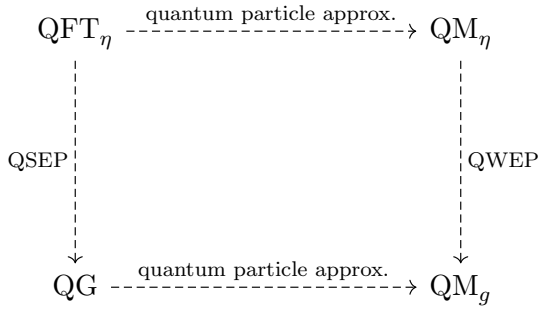
JCAP09(2019)057

classical Minkowski spacetime. A testimony of this fact is the quantum correction term for the metric (3.8), which features off-diagonal matrix elements of the metric operator $\hat{g}_{\mu\nu}$:

$$h_{\mu\nu} = 2\,\mathrm{Re}\left(\kappa\langle\Psi^{\perp}|\hat{g}_{\mu\nu}|\Psi\rangle\right) + \mathcal{O}(\eta)\,.$$

In this sense, due to the noncommutativity of the above diagram, one can argue that (within the discussed framework) *quantum gravity violates the weak equivalence principle*. Nevertheless, we would like to stress that our discussion regarding both strong and weak equivalence principles, based on the above prescription from subsection 4.1, is inherently *classical*. Indeed, in steps 1 and 2 which define the implementation of EP, one considers classical equations of motion. In our case, such definition suffices, as our entangled state (3.2) consists of a dominant and a sub-dominant term. Thus, we could expand our entangled equations (3.10) and (3.11) around the dominant classical terms, and discuss WEP in such a scenario. In fact, according to the definition of WEP, in general one can discuss its violation only *with respect to* some (perhaps unspecified, but assumed) classical spacetime metric. In our case, this role is played by the dominant classical metric $g_{\mu\nu}$.

In the more general case of superpositions of states which are more equally weighted, $\alpha \approx \beta$, and which consist of almost orthogonal states, $\langle\Psi|\tilde{\Psi}\rangle \approx 0$, one cannot single out a preferred classical metric, and therefore the classical definitions of SEP and WEP are inapplicable in this regime. Therefore, both equivalence principles ought to be extended to their respective *quantum* domains, denoted QSEP and QWEP respectively, in the sense of the following diagram:



Note that here all arrows are dashed, indicating the speculative nature of all these maps. Also, QM$_g$ represents a hypothetical theory of quantum particles coupled to a quantum gravitational field.

In this highly quantum regime ($\alpha \approx \beta$ and $\langle\Psi|\tilde{\Psi}\rangle \approx 0$), one could try to define the quantum weak equivalence principle (QWEP) in terms of the classical WEP, applied separately to each "branch" in the superposition. As long as the two "branches" $|\Psi\rangle$ and $|\tilde{\Psi}\rangle$ are themselves classical states, corresponding to the respective solutions of Einstein's equations, such a definition might seem suitable. Note that this approach is compatible with the notion of a superposed observer (see recent work [40] and the references therein). However, the formulation of the quantum strong and weak equivalence principles for the case of generic non-classical quantum states is an open question, outside of the scope of the current work.

Finally, the quantum version of the single pole approximation, called "quantum particle approximation" in the diagram above, is also not well defined — neither conceptually nor technically. Essentially, the whole diagram represents merely a speculation about the prescriptions which ought to map between the respective theories. In addition, like in the

JCAP09(2019)057

previous cases, the commutativity of the diagram (i.e., the violation of QWEP, given the validity of QSEP) would also be an open question. In some sense, the QSEP would represent a "true" equivalence principle, while QWEP would be a particle-like approximate image of QSEP. Being approximate, QWEP could possibly be violated in some cases, giving rise to noncommutativity of the diagram.

4.3 Universality, gravitational and inertial mass

In light of the results of section 3, in addition to the discussion of WEP violation, it is also important to discuss the status of the principle of universality, and the principle of equality between inertial and gravitational masses. In order to discuss them, it is instructive to study the Newtonian limit of the effective equation of motion (3.31), as follows.

We define the Newtonian limit in the standard way [39] — by assuming small spacetime curvature, nonrelativistic motion, and ignoring the backreaction of the particle on the background spacetime geometry. These approximations are implemented in the following way. First, ignoring the backreaction of the particle allows us to choose the dominant classical metric $g_{\mu\nu}$ as specified by the Newtonian line element

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + dx^2 + dy^2 + dz^2 \,, \tag{4.2}$$

where $x^\mu \equiv (t, x, y, z)$ are spacetime coordinates, M is the mass of the gravitational source, $r \equiv \sqrt{x^2 + y^2 + z^2}$, and $G \equiv l_p^2$ is Newton's gravitational constant. We will discuss the motion of a test-particle in this background, given by the effective equation of motion (3.31). Second, the assumption of nonrelativistic motion of the particle allows us to neglect its spacelike velocity,

$$u^k \equiv \frac{dz^k}{d\tau} \approx 0 \,,$$

leaving only the timelike component $u^0 \equiv dz^0/d\tau$ nonzero (the position of the particle $z^\mu(\tau)$ should not be confused with the label for the third spatial coordinate $z \equiv x^3$). Finally, the assumption of small spacetime curvature allows us to neglect all terms of order $\mathcal{O}(M^2)$ and higher.

Given this setup, one can easily calculate all nonzero Christoffel symbols corresponding to the dominant metric, obtaining:

$$\Gamma^0{}_{0k} = \Gamma^0{}_{k0} = \Gamma^k{}_{00} = \frac{GM}{r^3} x^k \,, \qquad k \in \{1, 2, 3\} \,.$$

One can then employ them to write the time and space components for the particle's effective equation of motion (3.31). Using (2.8) and (3.21), after some straightforward algebra, the time component of the equation of motion reduces to

$$\frac{d^2 z^0(\tau)}{d\tau^2} = 0 \,,$$

owing to the normalization condition $u^\mu u^\nu g_{\mu\nu} = -1$ and the presence of the orthogonal projector in (3.31). Using convenient initial conditions, this equation can be integrated to make an identification between the proper time τ and the time component of the particle's parametric equation of trajectory $x^\mu = z^\mu(\tau)$ as

$$t = z^0(\tau) = \tau \,,$$

JCAP09(2019)057

reflecting the notion of global universal time of Newtonian theory. Using this result, one can show that the space components of the particle's equation of motion obtain the following form (note that the spacelike indices can be raised and lowered at will, since the spatial part of the metric (4.2) is a unit matrix):

$$\frac{d^2 z^k}{d\tau^2} + \frac{GM}{r^3} z^k + \eta \left[\partial_0 h_{0k} - \frac{1}{2} \partial_k h_{00} - \frac{GM}{r^3} z^j h_{jk} \right] = 0 \,. \tag{4.3}$$

Note that here r has been evaluated at the position of the particle, $r = \sqrt{z^k z_k}$, and similarly for the gradients of h_{0k} and h_{00}. The first two terms in the equation come from the classical geodesic part ∇u^k in (3.31), while the third term is the quantum correction, coming from the effective force term $\eta u^\nu u^\sigma F^k_{\perp \nu\sigma}$.

The most important aspect of equation (4.3) is the similarity between the second term of the classical part and the final term of the quantum correction. The spacelike components h_{jk} can be separated into the trace and traceless part,

$$h_{jk} \equiv \frac{1}{3} h^i{}_i \, \delta_{jk} + \tilde{h}_{ij} \,, \qquad \tilde{h}^k{}_k \equiv 0 \,,$$

and the trace can be grouped together with the classical term, giving

$$\frac{d^2 z^k}{d\tau^2} + \frac{GM}{r^3} z^k \left(1 - \frac{1}{3} \eta h^i{}_i \right) + \eta \left[\partial_0 h_{0k} - \frac{1}{2} \partial_k h_{00} - \frac{GM}{r^3} z^j \tilde{h}_{jk} \right] = 0 \,. \tag{4.4}$$

Finally, multiplying the whole equation (4.4) with an arbitrary positive number, called the particle's *inertial mass* and denoted m_I, it takes the form of the Newton's second law of motion,

$$m_I \frac{d^2 z^k}{d\tau^2} = -m_I \left(1 - \frac{1}{3} \eta h^i{}_i \right) \frac{GM}{r^3} z^k - \eta m_I \left[\partial_0 h_{0k} - \frac{1}{2} \partial_k h_{00} - \frac{GM}{r^3} z^j \tilde{h}_{jk} \right] . \tag{4.5}$$

One can recognize two force terms on the right-hand side. The second term is of purely quantum origin, and represents the effective force acting on the particle ultimately due to the presence of the quantum state $|\tilde{\Psi}\rangle$ in (3.2). It has a non-Newtonian form, in the sense that none of its parts can be grouped together with the first force term, as was done with the trace part. The first force term, however, can be recognized as the classical Newton's gravitational force law, provided that one defines the ratio between the *gravitational mass* m_G and the *inertial mass* m_I of the particle as

$$\frac{m_G}{m_I} \equiv \left(1 - \frac{1}{3} \eta h^i{}_i \right) . \tag{4.6}$$

At this point we are ready to discuss the principles of universality and of the equality between gravitational and inertial masses. To begin with, it is obvious from (4.6) that the gravitational mass is equal to the inertial mass only up to a quantum correction term. This term contains the trace of spatial components of the metric interference tensor $h_{\mu\nu}$, defined by equation (3.8), from which we obtain:

$$h^i{}_i = 2\delta^{ij} \, \mathrm{Re} \left(\kappa \langle \Psi^\perp | \hat{g}_{ij} | \Psi \rangle \right) + \mathcal{O}(\eta) \,. \tag{4.7}$$

JCAP09(2019)057

It is crucial to notice that, in addition to the dependence of the off-diagonal matrix element of the metric operator, this expression also depends on the matter fields (which are present in $|\Psi\rangle$ and $|\Psi^{\perp}\rangle$), including the particle itself. Therefore, the term in the parentheses in (4.6) cannot be reabsorbed into the constants G and M, since these describe the external source of gravity which should remain independent of the properties of the test particle. Thus, the only possibility to cast the first force term in (4.5) into the form of the Newton's law of gravitation, is to define the ratio between the gravitational and the inertial mass as in (4.6). As a consequence, the principle of equality between gravitational and inertial masses is violated by the presence of the correction term coming from quantum gravity.

A similar argument can be made regarding the principle of universality. One may cancel away the inertial mass from the Newton's law (4.5), returning to (4.4) which describes the acceleration of the particle in the presence of an external gravitational field. Again, the presence of (4.7) in the classical gravitational acceleration term guarantees that this term depends not only on the external gravitational source, but also on the structure of the test particle itself. Moreover, the remaining quantum correction terms also depend on $h_{\mu\nu}$, and therefore they too carry information about the internal structure of the particle. In this sense, test particles described by different matter configurations may therefore display different accelerations, given the same background gravitational field. This means that the principle of universality is violated by the presence of the quantum gravity correction terms.

As a final comment, we should also note that m_I (and consequently m_G as well) is a completely free parameter in the Newtonian setup, and should be determined by the interactions of nongravitational type. In particular, the Newtonian framework does not allow us to connect m_I, m_G with the effective mass m of the particle, discussed in the context of (3.28) and (3.30). This is because the total rest-energy of a particle is an inherently relativistic concept, not defined in Newtonian mechanics. On the other hand, if one goes to the relativistic framework, the notions of inertial and gravitational masses become ill-defined, since gravitational interaction cannot be described anymore by a mere force law in the Newtonian sense. Therefore, the relationship between m on one side, and both m_I, m_G on the other side, remains undefined.

5 Conclusions

5.1 Summary of the results

In this paper, we have discussed the effective motion of a point particle within the framework of quantum gravity, in particular the case where both matter and gravity are in a quantum superposition of the Schrödinger cat type. In section 2 we gave a recapitulation of the results of the classical theory, introducing the multipole formalism framework and illustrating the derivation of the geodesic equation for the motion of a particle in GR. Section 3 was devoted to the generalization of these results to the realm of the full quantum gravity. In subsection 3.1 we introduced the abstract quantum gravity framework, discussed the model of the superposition of two classical states, and established the main assumptions for the derivation of the effective equation of motion. In subsection 3.2 we have analyzed in detail the quantum version of the equation for the covariant conservation of stress-energy tensor, which is a crucial ingredient in the derivation of the effective equation of motion. The explicit derivation of the equation of motion itself was then given in subsection 3.3, giving rise to the main results of the paper — the equation for the stress-energy kernel (3.28), the equation for the time-evolution of the particle's mass (3.30), and the effective equation of motion for

JCAP09(2019)057

the particle (3.31). Most importantly, the effective equation of motion turns out to contain a non-geodesic term, giving rise to an effective force acting on the particle, as a consequence of the interference terms between the two classical states of the gravity-matter system. The last subsection 3.4 discusses the self-consistency of the assumptions used in the above analysis, giving rise to the equation (3.35) for the error estimate of the single pole approximation scale.

In light of the nongeodesic motion established in section 3, it is important to discuss it in the context of the equivalence principle. This topic was taken up in section 4. After we have defined various flavors of the equivalence principle in subsection 4.1, the main analysis was presented in subsection 4.2, discussing a possible violation of (various forms of) the weak equivalence principle, as a consequence of the nongeodesic correction to the equation of motion (3.31). Also, given the inherently classical nature of the equivalence principle, we have also speculated on possible generalizations to the quantum realm, introducing the notions of the quantum strong and weak equivalence principles, albeit without giving explicit statements about their definitions. Finally, in subsection 4.3 we have discussed the notions of universality and equality between inertial and gravitational masses in the context of quantum gravity, by studying the Newtonian limit of the equation of motion (3.31). This analysis gave a clear interpretation that both universality and the equality between gravitational and inertial masses are violated in our context, corroborating the conclusions of the abstract analysis of the EP given in subsection 4.2.

5.2 Discussion of the results

By far the most interesting topic to discuss in the context of the equation of motion (3.31) is how to estimate the magnitude of the nongeodesic term. As far as the analysis of this paper goes, we can only say that this term is very small, given that it is proportional to η, which is in turn bounded from above by phenomenological argument that we do not observe superpositions of the gravitational field in nature. However, aside from this qualitative argument, in order to estimate the actual magnitude of the nongeodesic term one would need to go beyond the abstract quantum gravity formalism, and construct an explicit quantum gravity model coupled to matter fields, find some explicit kink solutions of the matter sector, and then calculate the overlap terms and the off-diagonal interference terms of the metric operator. Of course, any estimate obtained in such a way would be model-dependent. We consider this to be a feature of the abstract quantum gravity approach, since the magnitude of the nongeodesic term represents one way to operationally distinguish between different QG models. In other words, one could use equation (3.31) to experimentally test and compare these models, at least in principle. Probably the most obvious such test would employ equation (4.6) which relates the gravitational and inertial mass of the particle.

One result that was not discussed in detail is the nonconservation law for the effective mass of the particle, (3.30). However, it is not really surprising that the particle's total rest energy fails to be constant in the presence of gravity-matter entanglement. As (3.30) tells us, the nonconservation is actually a consequence of the additional effective force, which is itself a consequence of the quantum interference between two classical geometries and matter states. Nevertheless, it would indeed be interesting to study the mass nonconservation in more detail.

It is also important to discuss the generalization of our results from the case of the superposition of two classical states to many classical states. In particular, one could discuss the case where the state $|\tilde{\Psi}\rangle$ in (3.2) is not a single classical state, but a superposition of

JCAP09(2019)057

many classical states,

$$|\tilde{\Psi}\rangle = \sum_i \gamma_i |\Psi_i\rangle \,.$$

As long as we maintain the assumption that the fidelity $F(|\Psi\rangle, |\boldsymbol{\Psi}\rangle) \approx 1$, it is straightforward to see that all our results and conclusions still hold in the generic case. Therefore, there is no substantial difference in the analysis of a state which is a superposition of two classical states, compared to the analysis of a superposition of many classical states, as long as one of them is dominant while all others are sub-dominant. Note that in this case, even when β is finite and $\epsilon \to 0$, the role of the metrics generated by $|\Psi\rangle$ and $|\tilde{\Psi}\rangle$ cannot be exchanged anymore, as the latter generically does not satisfy Einstein's equations. This fixes the choice of $|\Psi\rangle$ as the dominant state. A detailed quantitative description is technically more complicated, but qualitatively all results will hold for both types of states.

5.3 Future lines of research

One of the main lines of future work would be to perform a similar analysis as was done in this paper, but keeping the η^2 terms. This would naturally include the sub-dominant effective metric and stress-energy (3.3), giving qualitatively new insight into the notion of quantum superpositions of two classical geometries. That analysis might provide clues about the properties of quantum gravity which could arguably hold even in the equal-weight superpositions of two classical states, defined by the choice $\alpha \approx \beta \approx 1/\sqrt{2}$ in (3.2).

Alternatively, one could repeat the analysis of this paper, but in a pole-dipole approximation. This would also lead to novel effects, one of which might be a coupling of various quantum interference terms to the spacetime curvature and the angular momentum of the particle, generalizing the classical pole-dipole equation of motion [10].

Also, given that the multipole formalism is also applicable to Riemann-Cartan spacetimes [19–22], the analysis of this paper could be generalized to include coupling of quantum interference terms to spacetime torsion and the spin of the particle.

Finally, one could further discuss a more general setup in which the off-diagonal terms in the covariant conservation equation (3.14) are not ignored, in the sense of going beyond the approximations (3.15) and (3.16).

In addition to all of the above, one important line of research would be to study possible connections to experiments. First, one should study the counterpart of the so-called *geodesic deviation equation*. Namely, in GR, the geodesic motion as such is not observable, as a consequence of the equivalence principle. As we have emphasized in subsection 4.1, the EP dictates that the only way to observe gravitational degrees of freedom is via *nonlocal measurements*, which are not encoded in the geodesic equation. Therefore, what one can actually observe is the change in the relative separation of two nearby geodesic trajectories, due to the tidal effects. This is in turn described by the geodesic deviation equation, which explicitly features the Riemann curvature tensor. In our case, the equation of motion (3.31) is not a geodesic, but is still local in character, in the sense that it does contain gravitational degrees of freedom at the given point, but still it does not combine gravitational degrees of freedom of two or more points. Thus, one ought to compare the trajectories of two nearby particles, both following a trajectory determined by (3.31). The equation governing the separation between two particles in such a setup would be a counterpart to the geodesic deviation equation of GR with a corresponding quantum correction term. It should be derived and studied in detail, in order to better understand what effects could be in principle directly experimentally observable.

JCAP09(2019)057

Second, one could also test our results by measuring the violation of the universality and of the equality of the gravitational and the inertial mass in the semiclassical Newtonian limit.

The above list of possible topics for further research is of course not exhaustive — one can probably study various additional aspects and topics related to this work, in particular giving more precise meaning to the notions of the quantum strong and weak equivalence principles.

Acknowledgments

NP acknowledges the support of SQIG (Security and Quantum Information Group), the Instituto de Telecomunicações (IT) Research Unit, Ref. UID/EEA/50008/2019, funded by Fundação para a Ciência e Tecnologia (FCT), project H2020 SPARTA, and the FCT projects Confident PTDC/EEI-CTP/4503/2014, QuantumMining POCI-01-0145-FEDER-031826 and Predict PTDC/CCI-CIF/29877/2017, supported by the European Regional Development Fund (FEDER) through the Competitiveness and Internationalization Operational Programme (COMPETE 2020), and by the Regional Operational Program of Lisbon, as well as the bilateral scientific cooperation between Portugal and Serbia through the project "Noise and measurement errors in multi-party quantum security protocols", no. 451-03-01765/2014-09/04 supported by the Foundation for Science and Technology (FCT), Portugal, and the Ministry of Education, Science and Technological Development of the Republic of Serbia. FCT grant CEECIND/04594/2017 is also acknowledged.

MV was supported by the project ON171031 of the Ministry of Education, Science and Technological Development of the Republic of Serbia, by the bilateral scientific cooperation between Portugal and Serbia through the project "Quantum Gravity and Quantum Integrable Models — 2015-2016", no. 451-03-01765/2014-09/24 supported by the Foundation for Science and Technology (FCT), Portugal, and the Ministry of Education, Science and Technological Development of the Republic of Serbia, by the bilateral scientific cooperation between Austria and Serbia through the project "Causality in Quantum Mechanics and Quantum Gravity — 2018-2019", no. 451-03-02141/2017-09/02 supported by the Austrian Academy of Sciences (ÖAW), Austria, and the Ministry of Education, Science and Technological Development of the Republic of Serbia, and by the "JESH Incoming 2016" grant of the Austrian Academy of Sciences (ÖAW).

NP and MV would also like to thank the Erwin Schrödinger International Institute for Mathematics and Physics (ESI Vienna), for the warm hospitality during the workshop "Quantum Physics and Gravity" (29 May – 30 June 2017) and partial support.

A Short review of the multipole formalism

In this appendix we give a short review of the multipole formalism, providing some basic motivation for its introduction and a few elementary properties. A more rigorous treatment and more details can be found in [25].

The multipole formalism revolves around the idea of expanding a function into a series of derivatives of the Dirac δ function, or δ series for short. Perhaps the easiest way to understand the δ series is to introduce it as a Fourier transform of a power series. For example, given a real-valued function $f(x)$, one can write it as a Fourier transform of $\tilde{f}(k)$ as

$$f(x) = \int_{\mathbb{R}} dk\, \tilde{f}(k)\, e^{ikx}\,. \tag{A.1}$$

JCAP09(2019)057

In principle, we can expand $\tilde{f}(k)$ into power series as

$$\tilde{f}(k) = \sum_{n=0}^{\infty} c_n k^n \,,$$

where c_n are some coefficients, substitute the expansion back into (A.1), and integrate term by term. Using the identity

$$k^n e^{ikx} = (-i)^n \frac{\partial^n}{\partial x^n} e^{ikx}$$

and the integral representation of the Dirac δ function

$$\delta(x) = \frac{1}{2\pi} \int_{\mathbb{R}} dk\, e^{ikx} \,,$$

we obtain

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} c_n \int_{\mathbb{R}} dk\, k^n\, e^{ikx} = \sum_{n=0}^{\infty} (-i)^n c_n \frac{d^n}{dx^n} \int_{\mathbb{R}} dk\, e^{ikx} \\ &= \sum_{n=0}^{\infty} 2\pi (-i)^n c_n \frac{d^n}{dx^n} \delta(x) \equiv \sum_{n=0}^{\infty} b_n \frac{d^n}{dx^n} \delta(x) \,. \end{aligned}$$

In the last step, we have merely renamed the coefficients in the expansion.

The above example is the most elementary construction of the δ series, providing some intuition. It is straightforward to see that one can generalize the procedure to perform the expansion around an arbitrary point z instead of zero, such that

$$f(x) = \sum_{n=0}^{\infty} b_n \frac{d^n}{dx^n} \delta(x - z) \,.$$

The coefficients b_n can be evaluated using the inverse formula,

$$b_n = \frac{(-1)^n}{n!} \int_{\mathbb{R}} dx\, (x - z)^n f(x) \,, \tag{A.2}$$

and are usually called *n-th order moments* of the function $f(x)$. From (A.2) one sees that the δ series is well defined for every function $f(x)$, which falls off to zero faster than any power of x at both infinities.

Let us study an instructive example. Let the function $f(x)$ be an ordinary Gaussian, peaked around the point x_0,

$$f(x) = \frac{1}{\sqrt{\pi}} e^{-(x - x_0)^2} \,.$$

One can evaluate the coefficients in the corresponding δ series using (A.2) to obtain:

$$b_n = \begin{cases} \displaystyle\sum_{k=0}^{n/2} \frac{(z - x_0)^{n-2k}}{4^k\, k!\, (n - 2k)!} & \text{for even } n \,, \\ \displaystyle -\sum_{k=0}^{(n-1)/2} \frac{(z - x_0)^{n-2k}}{4^k\, k!\, (n - 2k)!} & \text{for odd } n \,. \end{cases}$$

JCAP09(2019)057

It is important to note the following property — if the expansion point z does not coincide with the peak of the Gaussian, x_0, the magnitude of the coefficients b_n in general grows with n. For example, if $z - x_0 = 2$, we have

$$f(x) = \delta(x-z) - 2\frac{d}{dx}\delta(x-z) + \frac{9}{4}\frac{d^2}{dx^2}\delta(x-z) - \frac{11}{6}\frac{d^3}{dx^3}\delta(x-z) + \frac{115}{96}\frac{d^4}{dx^4}\delta(x-z) + \dots$$

However, if the expansion point coincides with the peak, $z - x_0 = 0$, the magnitude of the coefficients falls off as n grows:

$$f(x) = \delta(x-z) + \frac{1}{4}\frac{d^2}{dx^2}\delta(x-z) + \frac{1}{32}\frac{d^4}{dx^4}\delta(x-z) + \frac{1}{384}\frac{d^6}{dx^6}\delta(x-z) + \dots$$

From this simple example one can infer an important property of δ series — the coefficients b_n decrease as n grows, if the expansion point is near the peak of the function $f(x)$. Turning the argument around, if we require that the coefficients decrease with n,

$$|b_n| > |b_{n+1}|\,, \qquad \forall n \in \mathbb{N}_0\,,$$

this places *a restriction on the possible values of the expansion point* z. This is the crucial property of the δ series, and is being used to define the "position of the particle" which corresponds to a distribution of matter fields described by a localized function $f(x)$.

Also, assuming that the expansion point z has been chosen to be near the peak of the function, the decreasing nature of the coefficients b_n allows one to approximate the function $f(x)$ by a truncated series. This formalizes the intuitive idea that if one looks at some localized distribution of matter fields from "far away", it will look roughly as a point particle. The truncation point then quantifies the amount of "internal structure" that is known about $f(x)$. One can therefore study the function $f(x)$ at various approximation levels: the *single pole* approximation,

$$f(x) \sim b_0\delta(x-z)\,,$$

the *pole-dipole* approximation,

$$f(x) \sim b_0\delta(x-z) + b_1\frac{d}{dx}\delta(x-z)\,,$$

the *pole-dipole-quadrupole* approximation,

$$f(x) \sim b_0\delta(x-z) + b_1\frac{d}{dx}\delta(x-z) + b_2\frac{d^2}{dx^2}\delta(x-z)\,,$$

and so on.

It is completely straightforward to generalize the δ series to three (or more) dimensions, with the δ series of a function $f(\vec{x})$ around the point $\vec{z}$ defined as

$$f(\vec{x}) = \sum_{n=0}^{\infty} b_n^{i_1\dots i_n}\frac{\partial}{\partial x^{i_1}}\dots\frac{\partial}{\partial x^{i_n}}\delta^{(3)}(\vec{x}-\vec{z})\,. \tag{A.3}$$

Here the indices $i_1, \dots i_n$ take values 1, 2 and 3, and the inverse formula for the coefficients is

$$b_n^{i_1\dots i_n} = \frac{(-1)^n}{n!}\int_{\mathbb{R}^3} d^3x\,(x^{i_1} - z^{i_1})\dots(x^{i_n} - z^{i_n})f(\vec{x})\,. \tag{A.4}$$

JCAP09(2019)057

For example, in electrostatics, one can expand the charge density $\rho(\vec{x})$ localized around the point $\vec{z} = 0$ as

$$\rho(\vec{x}) = b_0 \delta^{(3)}(\vec{x}) + b_1^i \frac{\partial}{\partial x^i} \delta^{(3)}(\vec{x}) + \dots$$

According to (A.4), the coefficients are

$$b_0 = \int_{\mathbb{R}^3} d^3x \, \rho(\vec{x}) \equiv Q \,, \qquad \vec{b}_1 = -\int_{\mathbb{R}^3} d^3x \, \vec{x} \rho(\vec{x}) \equiv -\vec{p} \,,$$

where we recognize the total charge Q and the electrostatic dipole moment $\vec{p}$ of the source. Thus we have

$$\rho(\vec{x}) = Q\delta^{(3)}(\vec{x}) - \vec{p} \cdot \nabla \delta^{(3)}(\vec{x}) + \dots$$

Substituting the δ series expansion of $\rho(\vec{x})$ into the formula for the electrostatic potential,

$$\varphi(\vec{r}) = \int_{\mathbb{R}^3} d^3x \, \frac{\rho(\vec{x})}{|\vec{r} - \vec{x}|} \,,$$

and evaluating the integral, one obtains the familiar expression for the multipole expansion in electrostatics [41]:

$$\varphi(\vec{r}) = \frac{Q}{|\vec{r}|} + \frac{\vec{p} \cdot \vec{r}}{|\vec{r}|^3} + \dots$$

This example also illustrates what type of approximation is achieved with the truncation of the δ series.

Next we generalize to time-dependent functions. If the function $f(\vec{x})$ evolves in time, while remaining localized in space, one can expand it into δ series by choosing the most convenient reference point $z(t)$ at each moment of time,

$$f(\vec{x}, t) = \sum_{n=0}^{\infty} b^{i_1 \dots i_n}(t) \frac{\partial}{\partial x^{i_1}} \dots \frac{\partial}{\partial x^{i_n}} \delta^{(3)}(\vec{x} - \vec{z}(t)) \,, \tag{A.5}$$

where $t \in \mathbb{R}$ is a time variable, and the coefficients b are now time-dependent. Then one can introduce the proper time τ, and use the identity

$$\int_{\mathbb{R}} d\tau \, \delta(t - \tau) = 1$$

to rewrite (A.5) in a 4-dimensional manifestly Lorentz-invariant form

$$f(x) = \int_{\mathbb{R}} d\tau \sum_{n=0}^{\infty} b^{\mu_1 \dots \mu_n}(\tau) \partial_{\mu_1} \dots \partial_{\mu_n} \delta^{(4)}(x - z(\tau)) \,, \tag{A.6}$$

where we have relabeled $(\vec{x}, t) \equiv x$, introduced $z^0(\tau) = \tau$, used shorthand notation $\partial_\mu \equiv \partial/\partial x^\mu$, and defined $b^0 = b^{00} = b^{000} = \dots = 0$, since the time derivatives do not actually appear in (A.5). The introduction of these auxiliary timelike components of the b-coefficients, demanded by Lorentz invariance, gives rise to an additional gauge symmetry of the expansion coefficients, since only the "spatial" components carry nontrivial information about the function $f(x)$. This additional gauge symmetry is called *extra symmetry 1*, and is studied in detail in [25].

JCAP09(2019)057

Finally, one can make one more generalization, and introduce the notion of a δ series around a p-brane, a $(p+1)$-dimensional submanifold living in a D-dimensional spacetime manifold. Namely, we have seen that one can expand a function into a δ series around a point and around a one-dimensional line (equations (A.3) and (A.6), respectively). Generalizing in that direction, one can introduce the world-trajectory of a p-dimensional object through D-dimensional spacetime $\mathcal{M}$, with parametric equations $x^\mu = z^\mu(\xi^a)$ describing the trajectory as a $(p+1)$-dimensional submanifold $\Sigma \subset \mathcal{M}$. Here $\mu \in \{0, \dots, D-1\}$ and $a \in \{0, \dots p\}$, where x^μ are coordinates on $\mathcal{M}$ while ξ^a are intrinsic coordinates on Σ. Then, given a function $f(x)$ whose support is localized near the submanifold Σ, one can write its δ series expansion around Σ in a fully diffeomorphism- and reparametrization-invariant way as:

$$f(x) = \int_\Sigma d^{p+1}\xi \sqrt{-\gamma} \sum_{n=0}^{\infty} \nabla_{\mu_1} \dots \nabla_{\mu_n} \left[B^{\mu_1 \dots \mu_n}(\xi) \frac{\delta^{(D)}(x - z(\xi))}{\sqrt{-g}} \right] . \tag{A.7}$$

Here γ is the determinant of the induced metric $\gamma_{ab} = g_{\mu\nu} u_a^\mu u_b^\nu$ on Σ, where $g_{\mu\nu}$ is the metric on $\mathcal{M}$ and $u_a^\mu \equiv \partial z^\mu / \partial \xi^a$ are the tangent vectors of Σ. Note that, in order to ensure the correct tensorial behavior, the B-coefficients have been moved inside the action of the covariant derivatives. Namely, despite the fact that the covariant derivatives act with respect to x and B's do not depend on x, covariant derivatives still act nontrivially on B's with the connection terms. For similar reasons, the term $\sqrt{-g}$ has been introduced to combine with the δ function into a quantity which transforms as a scalar under diffeomorphisms. Its introduction amounts merely to a suitable redefinition of B's and does not modify the δ series in any nontrivial way.

The fully general δ series (A.7) has been studied in detail in [25]. For the purpose of the discussion given in the main text of this paper, we are interested in the case of a particle, i.e., a $(p=0)$-brane, moving along a 1-dimensional timelike curve $\mathcal{C}$ which is a submanifold of the $(D=4)$-dimensional spacetime $\mathcal{M}$. In this case, there is only one intrinsic coordinate on $\mathcal{C}$, denoted $\xi^0 \equiv \tau$, only one tangent vector

$$u_0^\mu \equiv \frac{\partial z^\mu(\xi)}{\partial \xi^0} = \frac{dz^\mu(\tau)}{d\tau} = u^\mu ,$$

while the induced metric tensor is a 1×1 matrix $\gamma_{00} = g_{\mu\nu} u_0^\mu u_0^\nu$. The parametrization of the curve $\mathcal{C}$ with the coordinate τ can be chosen to fix the reparametrization gauge symmetry via the gauge-fixing condition $\gamma_{00} = -1$, which is actually the natural normalization of the tangent vector, $g_{\mu\nu} u^\mu u^\nu = -1$. Finally, one can then apply the δ series expansion (A.7) to the stress-energy tensor $T^{\mu\nu}(x)$ of the matter fields as

$$T^{\mu\nu}(x) = \int_\mathcal{C} d\tau \sum_{n=0}^{\infty} \nabla_{\rho_1} \dots \nabla_{\rho_n} \left[B^{\mu\nu\rho_1 \dots \rho_n}(\tau) \frac{\delta^{(D)}(x - z(\tau))}{\sqrt{-g}} \right] .$$

Note that the coefficients B now carry two additional indices inherited from the stress-energy tensor. In the single pole approximation, one drops all terms in the sum except the $n=0$ term, truncating the series to the form

$$T^{\mu\nu}(x) = \int_\mathcal{C} d\tau \, B^{\mu\nu}(\tau) \frac{\delta^{(D)}(x - z(\tau))}{\sqrt{-g}} ,$$

as used in the main text.

JCAP09(2019)057

B Separable classical states

As mentioned in the main text, a recent study suggests that physical states of gravity and matter are generically entangled [1]. In this appendix, we analyze a simple, yet possibly intriguing, consequence of the assumption that the overall classical gravity-matter state can be approximated by (or indeed is) the product of the gravity and the matter classical states, $|\Psi\rangle = |g\rangle \otimes |\phi\rangle$, where $|g\rangle \in \mathcal{H}_G$ and $|\phi\rangle \in \mathcal{H}_M$ are classical states for the gravity and the matter sector, respectively (and analogously for $|\tilde{\Psi}\rangle$).

To begin with we introduce the overlaps as follows:

$$S_G \equiv \langle g|\tilde{g}\rangle\,, \qquad S_M \equiv \langle \phi|\tilde{\phi}\rangle\,, \qquad S \equiv \langle \Psi|\tilde{\Psi}\rangle = S_G S_M\,.$$

Note that, since in (3.2) only the relative phase between $|\Psi\rangle$ and $|\tilde{\Psi}\rangle$ is important, we can reabsorb the phases of the coefficients α and β into $|\Psi\rangle$ and $|\tilde{\Psi}\rangle$, respectively. In this way, we have $\alpha, \beta \in \mathbb{R}$, while only the overlap S between the two coherent states carries the information about the relative phase, and is therefore complex. Moreover, since S is a product between S_G and S_M, the phase of S can be distributed between S_G and S_M in an arbitrary way. A convenient choice is to have the phase in the matter sector, so that $S_G \in \mathbb{R}$ and $S_M \in \mathbb{C}$. Next, we can decompose $|\tilde{g}\rangle$ and $|\tilde{\phi}\rangle$ into parts proportional to and orthogonal to $|g\rangle$ and $|\phi\rangle$, respectively,

$$|\tilde{g}\rangle = S_G|g\rangle + \epsilon_G|g^\perp\rangle\,, \qquad |\tilde{\phi}\rangle = S_M|\phi\rangle + \epsilon_M|\phi^\perp\rangle\,, \tag{B.1}$$

where $\langle g|g^\perp\rangle \equiv 0$, $\langle \phi|\phi^\perp\rangle \equiv 0$, and

$$\epsilon_G \equiv \sqrt{1-(S_G)^2}\,, \qquad \epsilon_M \equiv \sqrt{1-|S_M|^2}\,.$$

Note that $\epsilon_G, \epsilon_M \in \mathbb{R}$. Additionally, one can use (B.1) to rewrite $|\tilde{\Psi}\rangle$ into the form

$$|\tilde{\Psi}\rangle = S|\Psi\rangle + \epsilon|\Psi^\perp\rangle\,,$$

where

$$\epsilon = \sqrt{\epsilon_M^2 + \epsilon_G^2 - \epsilon_M^2\epsilon_G^2}\,,$$

and

$$|\Psi^\perp\rangle = \frac{\epsilon_M S_G}{\epsilon}|g\rangle \otimes |\phi^\perp\rangle + \frac{\epsilon_G S_M}{\epsilon}|g^\perp\rangle \otimes |\phi\rangle + \frac{\epsilon_G\epsilon_M}{\epsilon}|g^\perp\rangle \otimes |\phi^\perp\rangle\,. \tag{B.2}$$

Note that in the cases when S_G and S_M are large (and consequently ϵ_G and ϵ_M are small), we can neglect the final term from (B.2), obtaining the Schmidt form of the "orthogonal correction" of the state $|\tilde{\Psi}\rangle$, with respect to $|\Psi\rangle$. It is interesting to observe that such a state is always necessarily entangled, as its entanglement entropy is always bigger than zero. In other words, to obtain a nearby classical product state of gravity and matter $|\tilde{\Psi}\rangle$, one has to perturb the original (classical product) state $|\Psi\rangle$ with an entangled state $|\Psi^\perp\rangle \approx \epsilon^{-1}\epsilon_M S_G|g\rangle \otimes |\phi^\perp\rangle + \epsilon^{-1}\epsilon_G S_M|g^\perp\rangle \otimes |\phi\rangle$.

C Phase of interference terms

In this appendix, we analyze the expressions for the expectation values of the metric and the stress-energy tensors in the entangled state (3.5), given by (3.6) and (3.7), respectively. We show that their third, interference, terms are generically different from zero, and thus

JCAP09(2019)057

contain non-trivial contributions linear in η. Since the two terms have the same form, we will consider the case of the metric operator only.

By writing $\kappa = |\kappa|e^{i\varphi_\kappa} = Fe^{i\varphi_\kappa}$ and $\langle\Psi^\perp|\hat{g}_{\mu\nu}|\Psi\rangle = |\langle\Psi^\perp|\hat{g}_{\mu\nu}|\Psi\rangle|e^{i\varphi_g}$, the third term of (3.6) has the form

$$2\eta\,\mathrm{Re}\left(\kappa\langle\Psi^\perp|\hat{g}_{\mu\nu}|\Psi\rangle\right) = 2\eta F|\langle\Psi^\perp|\hat{g}_{\mu\nu}|\Psi\rangle|\cos(\varphi_\kappa+\varphi_g)\,. \tag{C.1}$$

In case $\varphi_\kappa+\varphi_g = \pm\pi/2$, i.e., the interference term is zero, then for *any other generic* choice of $|\Psi'\rangle = e^{i\delta}|\Psi\rangle$, we have that $\varphi'_\kappa+\varphi'_g \neq \pm\pi/2$.

Indeed, changing $|\Psi\rangle \to |\Psi'\rangle = e^{i\delta}|\Psi\rangle$ induces the change of the other classical state

$$|\tilde{\Psi}\rangle = S|\Psi\rangle + \epsilon|\Psi^\perp\rangle \;\longrightarrow\; |\tilde{\Psi}'\rangle = S|\Psi'\rangle + \epsilon|\Psi^\perp\rangle = S'|\Psi\rangle + \epsilon|\Psi^\perp\rangle$$

with $S' = Se^{i\delta} = |S|e^{i(\varphi_s+\delta)}$ (where $S = |S|e^{i\varphi_s}$), but the orthogonal state $|\Psi^\perp\rangle$ *does not* change. Thus, the phase of the matrix element from (C.1) changes to $\varphi'_g = \varphi_g+\delta$. On the other hand, the phase of κ changes to $\varphi'_\kappa = \varphi_\kappa+\tilde{\delta}$ (note that $\tilde{\delta}$ is a function of δ, see below). Since

$$\begin{aligned}\kappa &= \alpha+\beta|S|e^{i\varphi_s} = |\kappa|e^{i\varphi_\kappa}\,,\\ \kappa' &= \alpha+\beta|S|e^{i(\varphi_s+\delta)} = |\kappa|e^{i\varphi'_\kappa}\,,\end{aligned}$$

it is obvious that for a generic choice of the parameters, i.e., in but a discrete number of points, we have $\tilde{\delta} = \varphi'_k-\varphi_k \neq -\delta$, obtaining $\varphi'_\kappa+\varphi'_g = (\varphi_s+\varphi_\kappa)+(\delta+\tilde{\delta}) = \pm\pi/2+(\delta+\tilde{\delta}) \neq \pm\pi/2$.

Thus, the linear correction to (3.6), and to (3.7) as well, is zero only for a discrete number of the relative phases between the classical states $|\Psi\rangle$ and $|\tilde{\Psi}\rangle$. Otherwise, it is generically non-trivial.

References

[1] N. Paunković and M. Vojinović, *Gauge protected entanglement between gravity and matter*, *Class. Quant. Grav.* **35** (2018) 185015 [arXiv:1702.07744] [INSPIRE].

[2] P. Chowdhury, D. Home, A.S. Majumdar, S.V. Mousavi, M.R. Mozaffari and S. Sinha, *Strong quantum violation of the gravitational weak equivalence principle by a non-Gaussian wave packet*, *Class. Quant. Grav.* **29** (2012) 025010 [arXiv:1107.5159] [INSPIRE].

[3] M. Zych and Č. Brukner, *Quantum formulation of the Einstein Equivalence Principle*, *Nature Phys.* **14** (2018) 1027 [arXiv:1502.00971] [INSPIRE].

[4] G. Rosi et al., *Quantum test of the equivalence principle for atoms in superpositions of internal energy eigenstates*, *Nature Commun.* **8** (2017) 5529 [arXiv:1704.02296] [INSPIRE].

[5] S. Longhi, *Equivalence principle and quantum mechanics: quantum simulation with entangled photons*, *Opt. Lett.* **43** (2018) 226 [arXiv:1712.02054].

[6] A. Einstein, L. Infeld and B. Hoffmann, *The Gravitational equations and the problem of motion*, *Annals Math.* **39** (1938) 65.

[7] M. Mathisson, *Neue mechanik materieller systemes*, *Acta Phys. Polon.* **6** (1937) 163 [INSPIRE].

[8] J. Lubánski, *Neue Bewegungsgleichungen materieller Systeme in Minkowskischer Welt*, *Acta Phys. Polon.* **6** (1937) 356.

[9] V.A. Fock, *Sur le mouvement des masses finie d'après la théorie de gravitation einsteinienne*, *J. Phys. USSR* **1** (1939) 81.

JCAP09(2019)057

[10] A. Papapetrou, *Spinning test particles in general relativity. 1.*, *Proc. Roy. Soc. Lond.* **A 209** (1951) 248 [INSPIRE].

[11] J. Weyssenhoff and A. Raabe, *Relativistic dynamics of spin-fluids and spin-particles*, *Acta Phys. Polon.* **9** (1947) 7 [INSPIRE].

[12] W. Tulczyjev, *Motion of multipole particles in general relativity theory*, *Acta Phys. Polon.* **18** (1959) 393.

[13] A.H. Taub, *Motion of Test Bodies in General Relativity*, *J. Math. Phys.* **5** (1964) 112.

[14] G. Dixon, *A covariant multipole formalism for extended test bodies in general relativity*, *Nuovo Cim.* **34** (1964) 317.

[15] G. Dixon, *Classical theory of charged particles with spin and the classical limit of the Dirac equation*, *Nuovo Cim.* **38** (1965) 1616.

[16] W.G. Dixon, *Dynamics of extended bodies in general relativity. I. Momentum and angular momentum*, *Proc. Roy. Soc. Lond.* **A 314** (1970) 499 [INSPIRE].

[17] W.G. Dixon, *Dynamics of extended bodies in general relativity. II. Moments of the charge-current vector*, *Proc. Roy. Soc. Lond.* **A 319** (1970) 509 [INSPIRE].

[18] G. Dixon, *The definition of multipole moments for extended bodies*, *Gen. Rel. Grav.* **4** (1973) 199.

[19] P.B. Yasskin and S.J. Stoeger, William R., *Propagation Equations for Test Bodies With Spin and Rotation in Theories of Gravity With Torsion*, *Phys. Rev.* **D 21** (1980) 2081 [INSPIRE].

[20] K. Nomura, T. Shirafuji and K. Hayashi, *Spinning test particles in space-time with torsion*, *Prog. Theor. Phys.* **86** (1991) 1239 [INSPIRE].

[21] K. Nomura, T. Shirafuji and K. Hayashi, *Semiclassical particles with arbitrary spin in the Riemann-Cartan space-time*, *Prog. Theor. Phys.* **87** (1992) 1275 [INSPIRE].

[22] M. Vasilić and M. Vojinović, *Zero-size objects in Riemann-Cartan spacetime*, *JHEP* **08** (2008) 104 [arXiv:0807.0596] [INSPIRE].

[23] C. Aragone and S. Deser, *String Dynamics from Energy Momentum Conservation*, *Nucl. Phys.* **B 92** (1975) 327 [INSPIRE].

[24] M. Vasilić and M. Vojinović, *Classical string in curved backgrounds*, *Phys. Rev.* **D 73** (2006) 124013 [gr-qc/0610014] [INSPIRE].

[25] M. Vasilić and M. Vojinović, *Classical spinning branes in curved backgrounds*, *JHEP* **07** (2007) 028 [arXiv:0707.3395] [INSPIRE].

[26] M. Vasilić and M. Vojinović, *Single-Pole Interaction of the Particle with the String*, *SIGMA* **4** (2008) 019 [arXiv:0802.1655] [INSPIRE].

[27] M. Vasilić and M. Vojinović, *Interaction of the particle with the string in pole-dipole approximation*, *Fortsch. Phys.* **56** (2008) 542 [arXiv:1010.1852] [INSPIRE].

[28] M. Vasilić and M. Vojinović, *Spinning branes in Riemann-Cartan spacetime*, *Phys. Rev.* **D 78** (2008) 104002 [arXiv:1010.1861] [INSPIRE].

[29] M. Vasilić and M. Vojinović, *Test membranes in Riemann-Cartan spacetimes*, *Phys. Rev.* **D 81** (2010) 024025 [arXiv:0812.4694] [INSPIRE].

[30] M. Vasilić, *Classical dynamics of brane-world extended objects*, *Phys. Rev.* **D 81** (2010) 124029 [INSPIRE].

[31] J. Armas, J. Camps, T. Harmark and N.A. Obers, *The Young Modulus of Black Strings and the Fine Structure of Blackfolds*, *JHEP* **02** (2012) 110 [arXiv:1110.4835] [INSPIRE].

[32] J. Armas, J. Gath and N.A. Obers, *Black Branes as Piezoelectrics*, *Phys. Rev. Lett.* **109** (2012) 241101 [arXiv:1209.2127] [INSPIRE].

JCAP09(2019)057

[33] J. Armas, J. Gath and N.A. Obers, *Electroelasticity of Charged Black Branes*, *JHEP* **10** (2013) 035 [arXiv:1307.0504] [INSPIRE].

[34] J. Armas, *How Fluids Bend: the Elastic Expansion for Higher-Dimensional Black Holes*, *JHEP* **09** (2013) 073 [arXiv:1304.7773] [INSPIRE].

[35] E. Okon and C. Callender, *Does Quantum Mechanics Clash with the Equivalence Principle — and Does it Matter?*, *Eur. J. Phil. Sci.* **1** (2011) 133 [arXiv:1008.5192] [INSPIRE].

[36] E. Di Casola, S. Liberati and S. Sonego, *Nonequivalence of equivalence principles*, *Am. J. Phys.* **83** (2015) 39 [arXiv:1310.7426] [INSPIRE].

[37] A. Accioly and R. Paszko, *Conflict Between the Classical Equivalence Principle and Quantum Mechanics*, *Adv. Studies Theor. Phys.* **3** (2009) 65.

[38] H.R. Brown and J. Read, *Clarifying possible misconceptions in the foundations of general relativity*, *Am. J. Phys.* **84** (2016) 327 [arXiv:1512.09253] [INSPIRE].

[39] C.W. Misner, K.S. Thorne and J.A. Wheeler, *Gravitation*, W.H. Freeman and Co., San Francisco, U.S.A. (1973).

[40] F. Giacomini, E. Castro-Ruiz and Č. Brukner, *Quantum mechanics and the covariance of physical laws in quantum reference frames*, *Nature Commun.* **10** (2019) 494 [arXiv:1712.07207] [INSPIRE].

[41] J.D. Jackson, *Classical Electrodynamics*, John Wiley and Sons Inc., New York, U.S.A. (1999).

JCAP09(2019)057

entropy

MDPI

Article

Quantum Contract Signing with Entangled Pairs

Preeti Yadav [1,2,*], Paulo Mateus [1,2], Nikola Paunković [1,2] and André Souto [1,3]

1 Instituto de Telecomunicações, 1049-001 Lisboa, Portugal
2 Departmento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
3 LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
* Correspondence: pri8.phy@gmail.com

Received: 11 July 2019; Accepted: 19 August 2019; Published: 22 August 2019



Abstract: We present a quantum scheme for signing contracts between two clients (Alice and Bob) using entangled states and the services of a third trusted party (Trent). The trusted party is only contacted for the initialization of the protocol, and possibly at the end, to verify clients' honesty and deliver signed certificates. The protocol is fair, i.e., the probability that a client, say Bob, can obtain a signed copy of the contract, while Alice cannot, can be made arbitrarily small, and scales as $N^{-1/2}$, where $4N$ is the total number of rounds (communications between the two clients) of the protocol. Thus, the protocol is optimistic, as cheating is not successful, and the clients rarely have to contact Trent to confirm their honesty by delivering the actual signed certificates of the contract. Unlike the previous protocol (Paunković et al., *Phys. Rev. A* **84**, 062331 (2011)), in the present proposal, a single client can obtain the signed contract alone, without the need for the other client's presence. When first contacting Trent, the clients do not have to agree upon a definitive contract. Moreover, even upon terminating the protocol, the clients do not reveal the actual contract to Trent. Finally, the protocol is based on the laws of physics, rather than on mathematical conjectures and the exchange of a large number of signed authenticated messages during the actual contract signing process. Therefore, it is abuse-free, as Alice and Bob cannot prove they are involved in the contract signing process.

Keywords: contract signing; quantum cryptography; quantum information

1. Introduction

Quantum cryptography traces back to late 1960s and early 1970s work on quantum money by Stephen Wiesner. While this work was published only a decade later, in 1983 [1], it had a significant impact on what usually is considered the birth of quantum cryptography, the seminal BB84 paper on quantum key distribution (QKD) [2]. Secure quantum communication offers higher, unconditional (i.e., information theoretic) security levels, as opposed to the computational security of classical cryptography. As a consequence, it became among the most prominent of the emerging quantum technologies, and QKD systems are currently available on the market, such as ID Quantique, QuintessenceLabs, etc. Other related protocols, such as quantum secure direct communication [3], have also been developed. Nevertheless, there is much beyond key distribution that quantum cryptography can offer, such as quantum secret sharing [4], quantum private query [5,6], and quantum secure distributed learning [7]. Secure multiparty computation (SMC) [8] presents another class of cryptographic protocols in which the privacy of users' data and inputs is protected. Instances of such schemes include private data mining and e-voting, to name a few. Recently, quantum solutions' bit commitment and oblivious transfer, cryptographic primitives that allow for execution of SMC, have been proposed [9]. In the current work, we present a quantum solution to the contract signing problem.

Contract signing [10] is a security protocol that falls within the group of the so-called commitment protocols [11–13]. In general, the protocol can be defined for an arbitrary number of parties (clients engaged in the protocol). For simplicity, we discuss the case of a two-party protocol, which can be straightforwardly generalized to an arbitrary number of participants.

The participants, usually referred to as Alice and Bob, have a common contract upon which they decide to commit, or not. The commitments are traditionally done by simple signatures: having a text of the contract with Bob's signature stamped on it, Alice can appeal to the authorities (the Judge), which in turn declares the document valid (i.e., binds the contract). In other words, having Bob's signature gives Alice the power to enforce the terms of the contract. Consequently, signing his name on a copy of the contract means that Bob commits to the contract. The aim of a contract signing protocol is that either both clients obtain each others' commitments or none of them do (the protocol is said to be fair). Further, if both clients follow the protocol correctly, both of them can obtain each others' commitments with certainty (the protocol is then said to be viable).

If only Alice has a copy with Bob's signature (i.e., only Bob is committed), she can later in time choose to either enforce, or not, the terms of the contract. Bob, however, has no power whatsoever: his future behavior is determined solely by Alice's decisions. For example, Alice may have a document with Bob's signature on it, declaring that he would buy a car from her, for a fixed amount of money. Knowing that only she has such a document, Alice can continue to negotiate the price of her car with other potential customers: in case she obtains a better offer, she is free to discard Bob's offer and thus is able to earn more money. Bob does not have such an option: if Alice does not obtain a better offer, she can always force Bob to buy the car from her, by showing to the authorities the contract signed by Bob. Having no proven commitment (signature) from Alice, Bob cannot enforce the contract himself and is thus unable to prevent Alice from such behavior, which puts him in an unfair situation.

Achieving fairness is trivial when clients meet up and simultaneously sign copies of the contract, thus both obtaining each others' commitments. Unfortunately, doing so when the clients are far apart, e.g., over the Internet, is difficult: indeed, sending his signed copy to Alice gives Bob no guarantee that he will obtain one from Alice; on the other hand, obtaining a signature from Alice before actually sending his gives Bob the advantage of having Alice's commitment without committing himself.

It has been shown [10,14] that the fairness of a contract signing protocol with spatially-distant clients can be achieved only by introducing a trusted third party, usually referred to as Trent, during the phase of exchanging clients' commitments. Trent's role is to receive clients' commitments and perform the exchange only upon obtaining signed copies of the contract from both clients. However simple and straightforward this solution may seem, it has a drawback, as Trent (in practice, a trusted agency accredited by the State that offers its time and resources for the exchange of money, e.g., public notaries) may be expensive. Therefore, the need for protocols using third parties as little as possible arises. Some contract signing protocols [15–17] do not require a trusted third party, but use a number of transmissions to send the pieces of signatures, or the partial information required to obtain the complete signature, in each message. Another possible way out is to design optimistic and/or probabilistic protocols.

In optimistic contract signing protocols [18], the exchange of commitments is, unless something goes wrong, executed solely by Alice and Bob. Only in case communication between the clients is interrupted (malfunction of the network, etc.), a trusted third party is involved [19]. In probabilistic protocols [20,21], by exchanging messages between each other, clients increase their probabilities to bind the contract. To be (probabilistically) fair, such protocols have to ensure that at each stage of the information exchange, the probabilities to bind the contract of both clients can be made arbitrarily close to each other (no client is significantly privileged). One such protocol is [21], for which the symmetry between the clients' positions is strengthened by the requirement that the joint probability that one client can bind the contract, while the other cannot, can be made smaller than any given $\varepsilon < 1$. Finally, there is both a probabilistically fair and optimistic solution, with an optimal number of exchanged

messages [20] for which even a stronger fairness condition is satisfied: the conditional probability that a client cannot bind the contract, when the other has already done so, can be made arbitrarily low.

Recently, a probabilistically fair and optimistic quantum protocol was presented in [22] (see a version using the simultaneous dense coding scheme in [23]). There, the trusted third party, Trent, was required to initiate the protocol and was contacted later only in case something went wrong. The protocol in [22] was also abuse-free [24], i.e., the clients cannot provide proofs of being involved in a contract signing procedure. Nevertheless, it has three important disadvantages: (i) Alice and Bob have to share the content of the contract with Trent; (ii) both clients have to be present in order to bind the contract, in case something goes wrong and Trent's services are required; and (iii) they have to agree upon the content of the contract before the protocol initialization. In this paper, we propose an improved version of the contract signing protocol where (i) the clients never disclose the content of the contract to Trent, (ii) only one client is needed to bind the contract, and (iii) the clients can decide upon the contract after they initially contact Trent.

Regarding Point (iii) from the previous paragraph, note that often, when parties initiate business negotiations, this does not result in making a deal formalized by a contract. Thus, involving Trent, who charges for his services, might often result in the waste of clients' resources, and Point (iii) might seem not to present a real advantage. Nevertheless, waiting for the last moment and contacting Trent only upon successful agreement and contract formulation might result in the system's failure due to possible communication bottlenecks. Imagine the following situation. Alice and Bob negotiate buying/selling a certain product, say a security system, knowing that on a given date in the future, a big company will announce a new model with its novel performances. Obviously, the price of the model used will highly depend on the information about the new one, and Alice and Bob will only upon learning the new piece of information decide upon the final contract. The problem is, many other users may choose to make similar business contracts in the same period of time, and if they all have to only then contact Trent, this might cause a communication bottleneck and the failure of the system. Thus, being able to contact Trent in advance and then, only later, "offline" (without contacting the trusted agency) exchange the commitments might be useful, especially in "more serious" business deals including larger amounts of money.

In Section 2, we begin with the description of the contract signing protocol with all the different phases explained in detail. In Section 3, we provide the security analysis of the protocol, together with relevant numerical results. Finally, in Section 5, we present the conclusions and discuss the contributions to the area.

2. The Proposed Protocol

In the quantum contract signing proposal [22], Trent sends two strings of qubits, one to Alice and another to Bob, such that each qubit is randomly prepared in one of the four BB84 states. The commitments are done by measuring one of the two observables on all qubits given to a client: in case of accepting the contract, measurement in the computational basis is performed, while choosing to reject it, one measures in the diagonal basis. Since the two bases are mutually unbiased, as a consequence of the Heisenberg uncertainty principle, it is impossible to learn both properties simultaneously. Thus, measurement outcomes of each client serve as certificates of the commitments. In order to achieve the fairness criterion (the two commitment choices have to be the same), the measurements are done in rounds, such that only one qubit per client is measured in each round and the outcomes are exchanged. Since, in addition to qubit strings, Trent also informs Alice of Bob's states (and vice versa), exchanging the outcomes allows clients to check each others' honesty. In our current proposal, instead of accompanying qubits with the classical information about other client's states, Trent exchanges entangled pairs. As a consequence of this change, in the current protocol, the choice of which out of the two observables clients measures in each round is different, as described below.

Consider two orthonormal qubit states $|0\rangle$ and $|1\rangle$ of the computational basis $\mathcal{B}_+$. The diagonal basis $\mathcal{B}_\times = \{|-\rangle, |+\rangle\}$ is given by $|\pm\rangle = \frac{1}{\sqrt{2}}(|1\rangle \pm |0\rangle)$. We also define two measurement observables:

$$\begin{aligned} \hat{O}_+ &= 1 \cdot |1\rangle\langle 1| + 0 \cdot |0\rangle\langle 0|, \\ \hat{O}_\times &= 1 \cdot |+\rangle\langle +| + 0 \cdot |-\rangle\langle -|. \end{aligned} \tag{1}$$

Let the bit string M be the contract upon which Alice and Bob agree. Let h be a publicly known hash function that they also agree to use. The value $h^* = h(M)$ will be the only information they provide to Trent about the contract M, when and if they contact him.

The protocol is described below in three parts: (i) initialization phase, Stage I: Alice and Bob contact the trusted party, Trent, who provides each of them with different secret classical information, to be used in the later phases of the protocol; Stage II: Trent prepares and distributes the states among the clients, to be used to sign the contract; (ii) exchange phase: the clients, using the states and the information provided by Trent, ping-pong the information needed to sign the contract; and (iii) binding phase: in this phase of the protocol, any one of the clients can contact Trent with his/her results in order to obtain an authorized document declaring the hash value $h^* = h(M)$ valid, which then validates the contract M.

Note that the exchange of classical information between Trent and the clients, in Stage I of the initialization phase, occurs over a private channel. During the rest of the protocol, the exchange of both quantum and classical information between Trent and the clients only needs to be authenticated. Classical authentication can be dealt with in a similar manner as in QKD, that is by assuming that a secret key is shared between Trent and the clients. How this key is exchanged depends on the level of security that we want to achieve: either an initial key is pre-shared (using a private channel), and then, it is enlarged using QKD (information-theoretical security); or a public key infrastructure is used (computational security). It is relatively easy to authenticate quantum information upon having classical authentication. It reduces to applying the cut-and-choose technique to verify the authenticity of quantum states exchanged, i.e., some random states are used by the clients and Trent to check whether the quantum channel was tampered with. This way, Trent discloses the description of the states over the authenticated classical channel, and the clients check if what they received is according to what was expected.

Below, we present the detailed description of the protocol.

Initialization phase: Stage I:

Parties involved: Alice, Bob, and Trent.

Input: Bit strings k_A and k_B of length $4N$ each and $2N$ randomly chosen indices from the set of indices of the $4N$ rounds of the protocol, $\mathcal{I} = \{1, 2, \ldots, 4N\}$.

Communication channel: Private classical channel between Trent and the clients.

Stage I of the initialization phase consists of the following steps:

1. Alice and Bob contact Trent for his services and inform him about the future time, at which Trent will begin Stage II of the initialization phase.
2. Trent provides Alice and Bob with randomly-generated bit strings k_A and k_B, respectively, of length $4N$ each. Alice and Bob prepare the strings $H_{\mathtt{A}} = h(M) \oplus k_A$ and $H_{\mathtt{B}} = h(M) \oplus k_B$ (bit-wise XOR), respectively. Note that in order to exchange the commitments to the whole string, we have $4N = |h(M)|$. We define the honest observables $\hat{H}_{\mathtt{A}_i}$ and $\hat{H}_{\mathtt{B}_i}$ for Alice and Bob, respectively, to be measured at each step i of the protocol ($i \in \mathcal{I}$) as:

$$\hat{H}_{\mathtt{A}_i} = \begin{cases} \hat{O}_+ & \text{if } H_{\mathtt{A}_i} = 0, \\ \hat{O}_\times & \text{if } H_{\mathtt{A}_i} = 1, \end{cases} \tag{2}$$

$$\hat{H}_{\mathtt{B}_i} = \begin{cases} \hat{O}_+ & \text{if } H_{\mathtt{B}_i} = 0, \\ \hat{O}_\times & \text{if } H_{\mathtt{B}_i} = 1, \end{cases} \tag{3}$$

where $H_{\mathtt{A}_i}$ is the i-th bit of the string $H_{\mathtt{A}}$ and analogously for Bob's string $H_{\mathtt{B}}$. Note that the secret keys k_A and k_B are used by the clients to hide their respective honest observables from each other. As it will be clear later on, the introduction of these keys prevents the scenario of a dishonest client, say, Bob, attacking the quantum channel between Trent and Alice by measuring the correct honest observables on Alice's qubits to obtain perfect correlations.

3. Trent provides Alice with a set of $2N$ randomly-chosen indices from the total 4N indices, $\mathcal{I}^{(A)} \subset \mathcal{I}$. Analogously, he randomly chooses $2N$ indices $\mathcal{I}^{(B)} \subset \mathcal{I}$ and sends them to Bob.
4. Trent provides Alice with $2N$ bits of Bob's secret string k_B, corresponding to the above-mentioned $2N$ indices from $\mathcal{I}^{(A)}$. Analogously, he sends the $2N$ bits to Bob from Alice's secret string k_A.

The initialization phase: Stage I ends with the following:

- Alice has a $4N$-long secret bit string k_A and a set $\mathcal{I}^{(A)} \subset \mathcal{I}$ of $2N$ randomly-chosen indices from $\mathcal{I}$. Alice also has k_{B_j} for all indices $j \in \mathcal{I}^{(A)}$.
- Bob has a $4N$-long secret bit string k_B and a set $\mathcal{I}^{(B)} \subset \mathcal{I}$ of $2N$ randomly-chosen indices from $\mathcal{I}$. Bob also has k_{A_j} for all indices $j \in \mathcal{I}^{(B)}$

Initialization phase: Stage II:

Parties involved: Alice, Bob, and Trent.
Input: $8N$ number of entangled pairs.
Communication channel: Authenticated classical and quantum channels between Trent and the clients.

The initialization phase: Stage II consists of the following steps (see Figure 1):

5. Trent produces two ordered sets, each consisting of $4N$ entangled pairs ($8N$ pairs in total). Each pair of particles is in the state $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle)$. He sends one particle from each pair of the first set to Alice, and from the second set to Bob, keeping the order of the pairs preserved. Let us denote the ordered set of $4N$ particles given to Alice by $\mathcal{A}$ and those given to Bob by $\mathcal{B}$. The two ordered sets kept by Trent, each consisting of $4N$ particles entangled with particles sent to Alice and Bob, are denoted by $\mathcal{T}^{(A)}$ and $\mathcal{T}^{(B)}$, respectively. We would like to note that the use of ordered sets was previously proposed in [3], later called the block transmission technique, crucial to quantum secure direct communication.
6. According to the set of $2N$ indices sent to Bob, $\mathcal{I}^{(B)}$, Trent divides $\mathcal{T}^{(A)}$ into two ordered subsets of $2N$ particles each: $\mathcal{T}_T^{(A)}$ and $\mathcal{T}_B^{(A)}$, with the $2N$ indices $j \in \mathcal{I}^{(B)}$ corresponding to the particles $\mathcal{T}_B^{(A)}$. Note that the original positions in $\mathcal{T}^{(A)}$ of each particle from $\mathcal{T}_T^{(A)}$ and $\mathcal{T}_B^{(A)}$ are preserved. In other words, for each particle from $\mathcal{T}_T^{(A)}$ and $\mathcal{T}_B^{(A)}$, Trent knows its position in $\mathcal{T}^{(A)}$ and, hence, with which particle in $\mathcal{A}$ it is entangled. The same is done with particles from $\mathcal{T}^{(B)}$, entangled with those in $\mathcal{B}$, obtaining $\mathcal{T}_T^{(B)}$ and $\mathcal{T}_A^{(B)}$.
7. Trent sends the ordered subsets $\mathcal{T}_B^{(A)}$ and $\mathcal{T}_A^{(B)}$ to Bob and Alice, respectively, each consisting of $2N$ particles. The particles in $\mathcal{T}_B^{(A)}$ and $\mathcal{T}_A^{(B)}$ are entangled with the corresponding particles ($2N$ of them) in the sets $\mathcal{A}$ and $\mathcal{B}$, given to Alice and Bob, respectively. Note that knowing the indices $\mathcal{I}^{(A)}$, Alice knows which particle from $\mathcal{T}_A^{(B)}$ is entangled with $2N$ of the particles in $\mathcal{B}$; and analogously for Bob. Trent keeps the subsets $\mathcal{T}_T^{(A)}$ and $\mathcal{T}_T^{(B)}$ to himself, to be used during the binding phase.

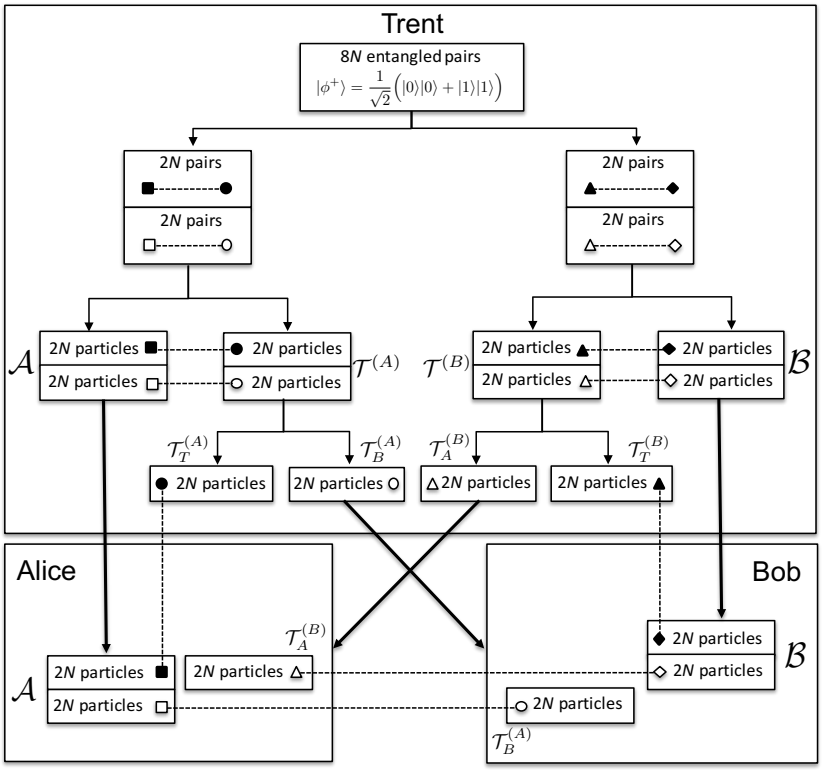


Figure 1. Initialization phase: Stage II: The thick arrows represent the transfer of particles from Trent to Alice and Bob. The dashed connections represent the entanglement between the respective particles. The big boxes represent the shielded private laboratories of Alice, Bob, and Trent.

The initialization phase: Stage II ends with the following:

- Alice has an ordered set $\mathcal{A}$ of $4N$ particles, entangled with $2N$ particles kept by Trent, $\mathcal{T}_T^{(A)}$, and additional $2N$ particles that are given to Bob, $\mathcal{T}_B^{(A)}$. She has another ordered set $\mathcal{T}_A^{(B)}$ of $2N$ particles, entangled with $2N$ particles given to Bob, chosen from $\mathcal{B}$ according to $\mathcal{I}^{(A)}$.
- Bob has an ordered set $\mathcal{B}$ of $4N$ particles, entangled with $2N$ particles kept by Trent, $\mathcal{T}_T^{(B)}$, and additional $2N$ particles that are given to Alice, $\mathcal{T}_A^{(B)}$. He has another ordered set $\mathcal{T}_B^{(A)}$ of $2N$ particles, entangled with $2N$ particles given to Alice, chosen from $\mathcal{A}$ according to $\mathcal{I}^{(B)}$.
- Trent keeps two ordered sets of $2N$ particles each, $\mathcal{T}_T^{(A)}$ and $\mathcal{T}_T^{(B)}$, entangled with $2N$ particles from $\mathcal{A}$ and $\mathcal{B}$, respectively.

Exchange phase:

Parties involved: Alice and Bob.
Input:

- The particles and indices Alice and Bob obtained at the end of the initialization phase.
- $h(M)$, the hash value of the contract M to be signed, obtained using publicly known function h.
- $H_{\mathtt{A}} = h(M) \oplus k_A$ and $H_{\mathtt{B}} = h(M) \oplus k_B$ for Alice and Bob, respectively.

Communication channel: Unauthenticated classical channel between the clients.

The exchange phase (see Figure 2) consists of $4N$ rounds. In each round, a client, say Alice, has a particle from $\mathcal{A}$, on which she measures $\hat{H}_{\mathtt{A}_i}$, with $i \in \mathcal{I}$, and sends the results to Bob. In addition to this, in $2N$ rounds labeled by $j \in \mathcal{I}^{(A)}$, Alice measures $\hat{H}_{\mathtt{B}_j}$ on the corresponding particles from $\mathcal{T}_A^{(B)}$. Note that Alice knows $k_{\mathtt{B}_j}$, and therefore, she knows $H_{\mathtt{B}_j} = h(M) \oplus k_{\mathtt{B}_j}$. Since Alice

knows that those particles are entangled with Bob, she uses her measurement outcomes to compare them with the results received from Bob, thus checking his behavior. Bob performs his measurements analogously. These two kinds of measurements are shown in Figure 3.

8. At the beginning of the exchange phase, Alice and Bob are in possession of $6N$ particles each. Alice has $4N$ particles denoted by $\mathcal{A}$ and $2N$ particles by $\mathcal{T}_A^{(B)}$, and analogously for Bob. On the $4N$ particles from $\mathcal{A}$, Alice measures her honest observable $\hat{H}_{\mathtt{A}_i}$, with $i = 1, \ldots, 4N$. Bob measures $\hat{H}_{\mathtt{B}_i}$ on his corresponding particles from $\mathcal{B}$. Their measurement outcomes form ordered sets of binary results $\mathcal{R}^{(A)}$ and $\mathcal{R}^{(B)}$, respectively, where:

$$\begin{aligned} \mathcal{R}^{(A)} &= \mathcal{R}_T^{(A)} \cup \mathcal{R}_B^{(A)}, \\ \mathcal{R}^{(B)} &= \mathcal{R}_T^{(B)} \cup \mathcal{R}_A^{(B)}. \end{aligned} \tag{4}$$

We use $\mathcal{R}_T^{(A)}$ to denote Alice's measurement results on the particles from $\mathcal{A}$ ($2N$ of them) that are entangled with $\mathcal{T}_T^{(A)}$ (kept by Trent) and $\mathcal{R}_B^{(A)}$ to denote her measurement results on the rest of the particles from $\mathcal{A}$ ($2N$ of them) entangled with $\mathcal{T}_B^{(A)}$ (given to Bob); and analogously for Bob's results, $\mathcal{R}_T^{(B)}$ and $\mathcal{R}_A^{(B)}$. They send these results to each other, one-by-one: if Alice starts first: in the i-th step of the exchange, she sends to Bob her result $\mathcal{R}_i^{(A)}$, then Bob sends to Alice his result $\mathcal{R}_i^{(B)}$, and so on.

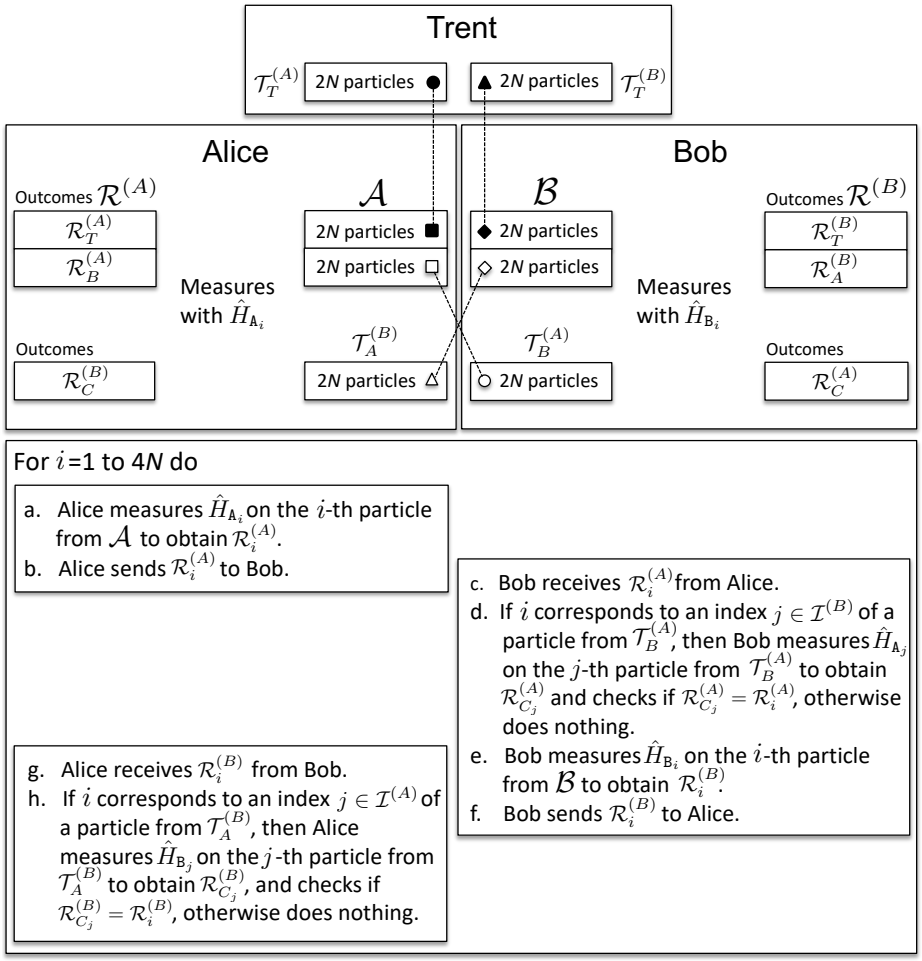


Figure 2. Exchange phase: the dashed connections represent the entanglement between the respective particles. The steps of the protocol during the exchange phase are described.

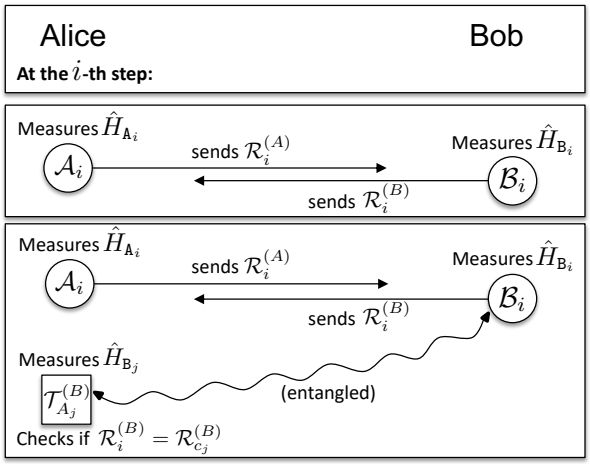


Figure 3. The two kinds of measurements a client, say, Alice, performs at a step i.

9. For each round $i \in \mathcal{I}$ for which there exists $j \in \mathcal{I}^{(A)}$, such that $i = j$, Alice measures $\hat{H}_{\mathrm{A}_j}$ on the corresponding particle from $\mathcal{T}_A^{(B)}$, to obtain $\mathcal{R}_{C_j}^{(B)}$ (see Figure 3). If Bob indeed measured his honest observable $\hat{H}_{\mathrm{B}_i}$, then his measurement outcome will match Alice's, $\mathcal{R}_i^{(B)} = \mathcal{R}_{C_j}^{(B)}$. In the presence of noise, Alice applies a statistical test to verify if Bob provided enough consistent results (see Appendix A.1.2). In case $i \neq j \in \mathcal{I}^{(A)}$, Alice uses Bob's result $\mathcal{R}_i^{(B)} \in \mathcal{R}_T^{(B)}$ for the optional binding phase, when Trent confronts Alice's information about $\mathcal{R}_i^{(B)}$ with his own measurement outcomes. The same is done by Bob upon receiving $\mathcal{R}_i^{(A)}$ from Alice. Then: (i) if all measurement outcomes, $\mathcal{R}_C^{(A/B)}$ ($2N$ of them for each client), are found to be consistent by the end of the communication at step $4N$, both clients will, during the binding phase, obtain with certainty the certified document from Trent that allows them to acquire a signed contract from the authorities (see the description below of the binding phase); (ii) if one of the clients suspects dishonest behavior, the communication is stopped, and they measure their honest observables on all remaining qubits and proceed to the binding phase.

The exchange phase ends with the following: If no cheating occurred, Alice and Bob both obtain their own, as well as each others' results, $\mathcal{R}^{(A)}$ and $\mathcal{R}^{(B)}$. In case the communication was interrupted at step m, a client, say, Alice, ends up with all of her own results $\mathcal{R}^{(A)}$ and those received from Bob by the step m (note that those do not necessarily need to be obtained by actually performing measurements on qubits).

Binding phase:

Parties involved: Trent and a client, say, Alice.
Input:

- The sets $\mathcal{T}_T^{(A)}$ and $\mathcal{T}_T^{(B)}$ of particles kept by Trent.
- The sets of Alice's measurement results, $\mathcal{R}^{(A)}$, and those sent to Alice by Bob, $\mathcal{R}^{(B)}$. Note that in the case of Bob's cheating, $\mathcal{R}^{(B)}$ might contain the wrong values, and in case the communication was interrupted at step m, it is only a partial set of results. For simplicity, we use the same symbol $\mathcal{R}^{(B)}$ for both sets of "honest", as well as "dishonest" results.
- $h^* = h(M)$, the hash value of the contract M to be signed, obtained using publicly-known function h.
- A publicly-known distribution $p(\alpha)$ to choose the acceptance rate α.

Communication channel: Private classical channel or in person.

The binding phase (see Figure 4) consists of the following step:

10. During the binding phase, a single client, say Alice, presents her results to Trent in order to bind the contract, to receive a certified document, signed by Trent's public key, declaring valid the hash value $h^* = h(M)$. Having such a certificate, Alice can appeal to the authorities to enforce the terms of the contract M: she presents the contract M and the signed document declaring the value h^* valid, so that the authorities can verify that indeed $h^* = h(M)$ (note that the function h is publicly known). As pointed out in the Introduction, Trent is an agency accredited by the State (e.g., public notaries). To bind the contract, Alice provides the string $h^* = h(M)$ to Trent and presents her results $\mathcal{R}^{(A)}$, as well as those obtained from Bob, $\mathcal{R}^{(B)}$ (if the protocol was interrupted before its completion, Alice will guess the rest of Bob's outcomes; see Appendix for details). Knowing h^*, Trent computes $H_{\mathtt{A}_j} = h^* \oplus k_{\mathtt{A}_j}$ and $H_{\mathtt{B}_j} = h^* \oplus k_{\mathtt{B}_j}$. Trent thus measures the honest observables $\hat{H}_{\mathtt{A}_i}$ on N randomly-chosen qubits from the subset $\mathcal{T}_T^{(A)}$ and $\hat{H}_{\mathtt{B}_i}$ on another N randomly-chosen qubits from $\mathcal{T}_T^{(B)}$ (the other $2N$ particles are kept for binding the contract for Bob, if requested). He also chooses independently at random $\alpha > 1/2$, according to some publicly-known distribution $p(\alpha)$. Trent will give to Alice a certified document declaring the hash value h^* valid ("bind the contract for Alice") if the results $\mathcal{R}^{(A)}$ and $\mathcal{R}^{(B)}$ satisfy the following two conditions:

10a. at least a fraction α of N Alice's results from $\mathcal{R}_T^{(A)}$ is equal to Trent's results on the corresponding entangled N particles from $\mathcal{T}_T^{(A)}$, and

10b. at least a fraction α of N Bob's results from $\mathcal{R}_T^{(B)}$ is equal to Trent's results on the corresponding entangled N particles from $\mathcal{T}_T^{(B)}$.

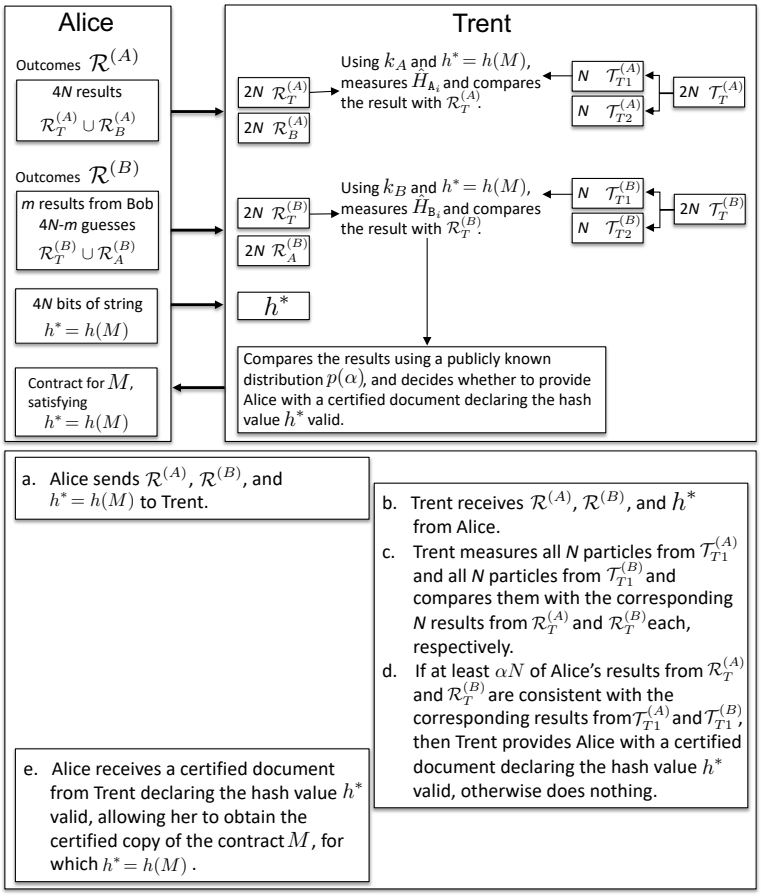


Figure 4. Binding phase: the thick arrows represent the transfer of measurement outcomes from Alice to Trent. The steps for the binding phase are described.

The binding phase ends with the following: In case Trent finds Alice's results consistent with his, she receives an authorized document from him declaring the hash value h^* valid, which then allows her to obtain the certified copy of the contract M, for which $h^* = h(M)$.

3. Security Analysis

In our protocol, a cryptographic hash function h is used to map contract M to a bit string of fixed size $4N = |h(M)|$. Had Trent possessed an infinite computational power, he would be able to find the collisions, among which one message would be the contract M. Nevertheless, the problem of finding collisions for existing cryptographic hash functions (such as SHA256 and others) is not based on any mathematical or number theoretical conjecture, such as the hardness of factoring, but on the fact that the hashing function is highly irregular and non-linear. Its security is at the same level of symmetric cryptography (such as AES), which is assumed to be beyond the capacity of quantum technologies to attack, and moreover, AES is actually used in current commercial QKD services. Furthermore, note that at the time this paper was written, it was not yet found a single collision for SHA256 of two meaningless texts, and so, it is unforeseeable to find collisions for a given fixed text. Google used more than 9×10^{18} hashes to find a meaningless collision of SHA1 [25], and SHA256 is considerably harder than SHA1. Finally, it is worthwhile noticing that the assumption that there exists an unbreakable hash function, the so-called random oracle model, is quite common, even when quantum information and computation is available [26]. In addition, having such computational power would also allow a cheating client (say, Bob) to find collisions as well, thus potentially giving him the opportunity to bind a different contract M', for which $h(M') = h(M)$. Nevertheless, given a particular hash function h, it is negligible that other collisions different from the contract M would still represent meaningful contracts, let alone contracts that would be favorable to Bob.

Let us define the following probabilities for Alice to pass the above tests (a) and (b), in case the communication was interrupted at step m,

$$\begin{aligned} P_{BTH}(m;\alpha) &\quad - \quad \text{Probability that Bob passes Trent's test on his own qubits,} \\ P_{BTA}(m;\alpha) &\quad - \quad \text{Probability that Bob passes Trent's test on Alice's qubits,} \end{aligned}$$

$P_{ATH}(m;\alpha)$ and $P_{ATB}(m;\alpha)$ can be analogously defined as Alice's probabilities to pass Tests (a) and (b). Additionally, we define Bob's probability to pass Alice's verification test on the results $\mathcal{R}_i^{(A)} \in \mathcal{R}_B^{(A)}$ (see Step (5) of Section 2) received from Alice by step m, as:

$$P_{BAS}(m) \quad - \quad \text{Probability that Bob passes Alice's test on their shared qubits,}$$

and analogously $P_{ABS}(m)$ for Alice.

It is easy to verify that, in the noiseless scenario, the protocol is optimistic. If both clients follow the protocol honestly till the end, both of them are able to enforce the contract: Alice will have all the consistent results for her own, as well as Bob's qubits, allowing her to bind the contract with probability one (the same happens for Bob).

To analyze the probabilistic fairness quantitatively, we introduce the so-called probability to cheat, along the lines of the similar quantity analyzed in [22]. By $P_{bind}^{\mathcal{B}}(m;\alpha) = P_{BTH}(m;\alpha) \cdot P_{BTA}(m;\alpha)$, we denote the probability that Bob passes Trent's tests and can thus bind the contract, if the communication is interrupted at step m of the protocol, for a given choice of α; and analogously for Alice. To reach step m, both clients have to pass each others' verification, which is given by the probability $P(m) = P_{BAS}(m) \cdot P_{ABS}(m)$. Bob's probability to cheat at step m, for a given α, is defined as the probability that he can bind the contract, while Alice cannot, multiplied by the probability to reach step m:

$$P_{ch}^{\mathcal{B}}(m;\alpha) = P(m) \cdot P_{bind}^{\mathcal{B}}(m;\alpha) \left[1 - P_{bind}^{\mathcal{A}}(m;\alpha)\right]. \tag{5}$$

Note that the above probabilities also depend on the particular distribution of entangled pairs, denoted as "configuration $\mathcal{L}$", given by probability $q(\mathcal{L})$, and in the case of a dishonest client, the cheating strategy. Furthermore, both the above, as well as any probability evaluated (with the exception of $p(\alpha)$) depend on N; therefore, we omit writing it, as it is implicitly assumed. Nevertheless, the dependence on configuration $\mathcal{L}$ is relevant in calculations, and below, we analyze it in detail.

As prescribed by the protocol, Trent gives $6N$ qubits to Alice: $4N$ qubits from $\mathcal{A}$ and $2N$ from $\mathcal{T}_A^{(B)}$ (see Figure 5), together with their relative positions. Analogously, Bob receives $6N$ qubits from $\mathcal{B}$ and $\mathcal{T}_B^{(A)}$. We assume that all the classical communications between Trent and clients are private and authenticated, based on, say, pre-shared symmetric key schemes. After the communication has stopped at step m, out of the $4N$ qubits to measure from $\mathcal{A}$ and $\mathcal{B}$, Alice and Bob will be left with $\ell^{(A)}$ and $\ell^{(B)}$ unmeasured qubits, respectively, with $\ell^{(A)} = \ell^{(B)} = 4N - m$. Note that, among the $\ell^{(A)}$ and $\ell^{(B)}$ qubits, not all will be used by Trent to bind the contract for Alice and Bob. In fact, the qubits that are entangled between Alice and Bob are irrelevant for their binding probabilities. They are used by the parties to check each others' honesty. Let then $\ell^{(A)}$ be decomposed into $\ell_T^{(A)}$ and $\ell_B^{(A)}$, the qubits entangled with those held by Trent and by Bob, respectively. Analogously, let $\ell^{(B)}$ be decomposed into $\ell_T^{(B)}$ and $\ell_A^{(B)}$. Therefore:

$$\begin{aligned} \text{Alice} &: \quad (2N+2N)-m \;=\; \ell^{(A)} \;=\; \ell_T^{(A)} + \ell_B^{(A)} \\ \text{Bob} &: \quad (2N+2N)-m \;=\; \ell^{(B)} \;=\; \ell_T^{(B)} + \ell_A^{(B)}. \end{aligned} \tag{6}$$

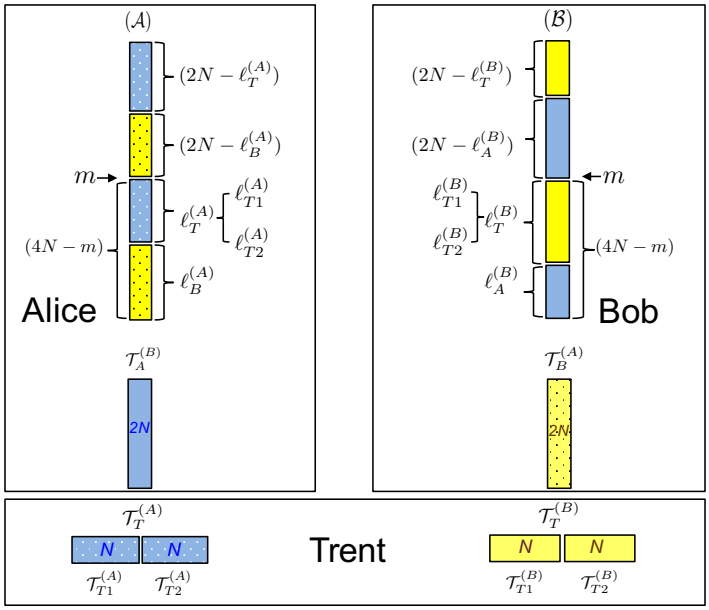


Figure 5. Decomposition of $\ell^{(A)}$ into $\ell_T^{(A)}$ and $\ell_B^{(A)}$ and decomposition of $\ell^{(B)}$ into $\ell_T^{(B)}$ and $\ell_A^{(B)}$. The same colors and patterns represent entanglement between the respective particles of different parties.

In order to bind the contract, a client, say Bob, has to present his own measurement results, as well as those obtained from Alice. Then, Trent checks if they are correlated with those obtained on qubits in his possession. Unlike the previous proposal [22], in which both clients had to be present and show their results to Trent in order to both obtain signed contracts during the binding phase, in the current protocol, Bob does not need Alice to be summoned in order to bind the contract (and vice versa). Since the protocol should be symmetric to both clients, it should allow that they both, separately, are able to bind the contract. For this reason, when binding the contract to, say Bob, Trent does not check all of his results from $\mathcal{T}_T^{(A)}$ and $\mathcal{T}_T^{(B)}$ for qubits entangled with $\mathcal{A}$ and $\mathcal{B}$, respectively. Note that to check Bob's results, Trent has to measure the honest observables $\hat{H}_{\mathtt{A}_i}$ and $\hat{H}_{\mathtt{B}_i}$ on his qubits, he obtains

using the $h^* = h(M)$ provided by Bob. Therefore, if both clients were using the same sets of qubits (entangled with those in Trent's possession) to bind the contract separately, a dishonest Bob would have a trivial successful cheating strategy. He measures his honest observables given by the mutually-agreed contract M, which allows him to bind that contract. Nevertheless, in case he later decides not to comply with it, he simply provides Trent with a random $h' \neq h(M)$. As a consequence, Trent's results will be uncorrelated with both Bob's, as well as Alice's results, i.e., neither client would be able to bind the contract M. This is precisely the reason for checking only N out of $2N$ qubits from $\mathcal{T}_T^{(A)}$ and $\mathcal{T}_T^{(B)}$, each.

Thus, Trent's qubits are each divided into two equal subsets of the same size, $\mathcal{T}_T^{(A)} = \mathcal{T}_{T1}^{(A)} \cup \mathcal{T}_{T2}^{(A)}$ and $\mathcal{T}_T^{(B)} = \mathcal{T}_{T1}^{(B)} \cup \mathcal{T}_{T2}^{(B)}$: the sets with the $T1$ subscript are used for binding the contract to Alice, while those with $T2$ for Bob. Consequently, we have $\ell_T^{(A)} = \ell_{T1}^{(A)} + \ell_{T2}^{(A)}$ and $\ell_T^{(B)} = \ell_{T1}^{(B)} + \ell_{T2}^{(B)}$ (see Figure 5).

The overall configuration $\mathcal{L}$ of the entangled pairs distributed between Alice, Bob, and Trent is given by six numbers,

$$\begin{aligned} \mathcal{L} &= \left\{ \ell_B^{(A)},\ \ell_{T1}^{(A)},\ \ell_{T2}^{(A)},\ \ell_A^{(B)},\ \ell_{T1}^{(B)},\ \ell_{T2}^{(B)} \right\}, \\ \text{with} \qquad & q(\mathcal{L}) \leq q(\ell_B^{(A)}) \cdot q(\ell_{T1}^{(A)}) \cdot q(\ell_{T2}^{(A)}) \cdot q(\ell_A^{(B)}) \cdot q(\ell_{T1}^{(B)}) \cdot q(\ell_{T2}^{(B)}). \end{aligned} \tag{7}$$

Therefore, Bob's probability to cheat, given by Equation (5), now written with the explicit dependence on the configuration $\mathcal{L}$, is:

$$P_{ch}^{\mathcal{B}}(m;\alpha|\mathcal{L}) = P(m|\ell_B^{(A)},\ell_A^{(B)}) \cdot P_{bind}^{\mathcal{B}}(m;\alpha|\ell_{T2}^{(A)},\ell_{T2}^{(B)}) \cdot \left[1 - P_{bind}^{\mathcal{A}}(m;\alpha|\ell_{T1}^{(A)},\ell_{T1}^{(B)})\right]. \tag{8}$$

Averaging the "constituent" probabilities from the above equation over their respective configurations from $\mathcal{L}$ gives:

$$\begin{aligned} P(m) &= \langle P_{ABS}(m|\ell_B^{(A)})\rangle_{\ell_B^{(A)}} \cdot \langle P_{BAS}(m|\ell_A^{(B)})\rangle_{\ell_A^{(B)}}, \\ P_{bind}^{\mathcal{B}}(m;\alpha) &= \langle P_{BTH}(m;\alpha|\ell_{T2}^{(B)})\rangle_{\ell_{T2}^{(B)}} \cdot \langle P_{BTA}(m;\alpha|\ell_{T2}^{(A)})\rangle_{\ell_{T2}^{(A)}}, \\ P_{bind}^{\mathcal{A}}(m;\alpha) &= \langle P_{ATH}(m;\alpha|\ell_{T1}^{(A)})\rangle_{\ell_{T1}^{(A)}} \cdot \langle P_{ATB}(m;\alpha|\ell_{T1}^{(B)})\rangle_{\ell_{T1}^{(B)}}, \end{aligned} \tag{9}$$

where $\langle A\rangle_\ell$ represents the expectation value of $A(\ell)$ over the values of ℓ. To simplify notation, in the following, we will use $P_{ABS}(m) = \langle P_{ABS}(m|\ell_B^{(A)})\rangle_{\ell_B^{(A)}}$ and $P_{BTH}(m;\alpha) = \langle P_{BTH}(m;\alpha|\ell_{T2}^{(B)})\rangle_{\ell_{T2}^{(B)}}$, and analogously for the other four probabilities from the right-hand sides of the above three equations.

Hence, with Bob's probability to cheat, averaged over all configurations $\mathcal{L}$,

$$\begin{aligned} P_{ch}^{\mathcal{B}}(m;\alpha) &= \sum_{\mathcal{L}} q(\mathcal{L})\ P_{ch}^{\mathcal{B}}(m;\alpha|\mathcal{L}) \\ &\leq P(m) \cdot P_{bind}^{\mathcal{B}}(m;\alpha) \left[1 - P_{bind}^{\mathcal{A}}(m;\alpha)\right], \end{aligned} \tag{10}$$

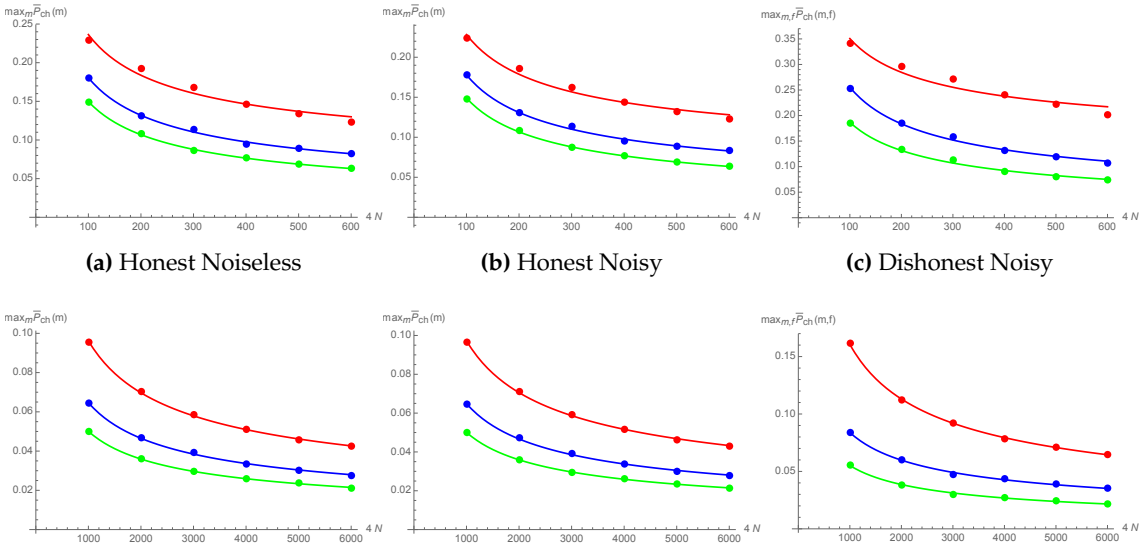
we have the expected probability to cheat as:

$$\bar{P}_{ch}^{\mathcal{B}}(m) = \int p(\alpha)\ P_{ch}^{\mathcal{B}}(m;\alpha)\ \mathrm{d}\alpha. \tag{11}$$

For honest clients that follow the protocol, the above probability is determined by the steps prescribed by the protocol (the "honest strategy"). In case a client, say Bob, does not follow the protocol, the above probability depends on the "cheating strategy" of a dishonest client. It turns out (see Appendix A.2) that the quantum part of the honest and the optimal cheating strategies is the same, i.e., the (quantum) measurements performed by a cheating Bob are the same as that of an honest one, given by his honest observables $\hat{H}_{\mathrm{B}_i}$. In other words, the best a cheating Bob can do is to send to Alice

the wrong results determined by a frequency f. This is a consequence of the fact that Bob does not know which of the qubits given to him are used to bind the contract by Trent and which to check his honesty by Alice (for details, see Appendix A.2).

In Appendix A.1, we derive the explicit expressions for the expected probability to cheat (11) for honest clients that follow the protocol, in the ideal noiseless case (Appendix A.1.1), as well as for noisy environments (Appendix A.1.2), thus showing the soundness of the protocol. In Figure 6a,b,d,e, we present the values of the maximal expected probability to cheat as a function of the total number of photons for the values of $4N$ up to 6000. In both cases (as well as for the case of a cheating client discussed below), the results are obtained for the uniform $p(\alpha)$ on the intervals $[0.9, 0.99]$, $[0.8, 0.99]$, and $[0.7, 0.99]$, and with a noise parameter $\kappa = 0.05$. The observed dependence $\max_m \bar{P}_{ch}^{\mathcal{A}/\mathcal{B}}(m) \propto N^{-1/2}$ is confirmed by the proof of the asymptotic behavior, $P_{ch}(m;\alpha) \in \mathcal{O}(N^{-1/2})$ (see Theorem A1 from Appendix A.1.1).



(a) Honest Noiseless **(b)** Honest Noisy **(c)** Dishonest Noisy

(d) Honest Noiseless **(e)** Honest Noisy **(f)** Dishonest Noisy

Figure 6. (**a**,**d**) correspond to the honest noiseless scenario, where the maximal expected probability to cheat, $\max_m \bar{P}_{ch}(m)$, is plotted against the total number of rounds between Alice and Bob, $4N$. (**b**,**e**) correspond to the case of honest clients in a noisy channel with a noise parameter of $\kappa = 0.05$. (**c**,**f**) correspond to the realistic case of a dishonest client with noisy channels, where the optimal cheating strategy depends on a parameter f. In all three cases, the red, blue, and green curves are obtained for the uniform $p(\alpha)$ on the intervals $[0.9, 0.99]$, $[0.8, 0.99]$, and $[0.7, 0.99]$, respectively.

In Appendix A.2, we evaluated the corresponding probabilities for the case of a cheating client who deviates from the protocol, in the presence of noise. In Figure 6c,f, we plot the values of the maximal expected probability to cheat against $4N$, for the optimal cheating strategy, showing the same dependence $\propto N^{-1/2}$. Further, in Appendices A.1 and A.2, we analyze the decrease of the expected probabilities to cheat in case the cheating strategy deviates from the optimal values of m or f.

The results presented in Figure 6 are obtained for a fixed value of the noise parameter κ. By increasing the noise, more and more "wrong" results are going to be obtained, such that even honest participants will either interrupt the communication during the exchange phase or will not be able to bind the contract with Trent. The figure of merit here is the final average probability to bind obtained for $m = 4N$, given by Equations (A26)–(A28) from Appendix A.1.2 (note that, if upon exchanging all the messages, clients have high enough probability to bind the contract, then they would also be able to pass each others' tests during the whole exchange phase with equally high probability).

Thus, for a fixed number of rounds, $4N$, range of α, and its probability distribution $p(\alpha)$, one can straightforwardly obtain the threshold values for the noise parameter κ, both for the honest noisy,

as well as for dishonest noisy cases (note that such threshold values depend on predetermined security level, i.e., the lower bound for the binding probability set by the users).

While such quantitative numbers can straightforwardly be obtained using the analysis presented in the paper, they would not be very informative. Namely, what one should do is to, given the noise level (given κ), determined by the actual implementation setup, optimize the rest of the relevant parameters (N, $p(\alpha)$ and its range). While conceptually, this is possible to do, it is clearly exceptionally demanding regarding the computational resources (note that in our analysis, we probed only three α ranges and only the simplest uniform distribution). Our paper is more of a proof of a concept, rather than the final analysis, which, as mentioned, is strongly implementation dependent.

Therefore, considering our limited computational power and the fact that the presented threshold values for κ would probably differ from the ones to be obtained by optimizing the parameters of actual implementations, we decided to omit such numerical analysis. Note that the above discussion also applies for the cheating probability: to obtain the optimized cheating probability levels, one should vary all the relevant parameters. Nevertheless, while it is obvious that the binding probability will decrease as the noise increases, it was not at all obvious that it is even possible to establish upper bounds for the cheating probability, such that it can be made arbitrarily low.

The techniques presented in our paper use the "brute force" numerical approach in obtaining the final quantitative results (with the exception of our analytic proof of the asymptotic behavior given in Theorem A1 from Appendix A.1.1), which do not allow for drawing qualitative insights. Developing more closed analytic expressions for the final binding and cheating probabilities that can be analyzed beyond the final numerical values would be an interesting topic of future research.

4. Discussion

Like the previous version of the quantum contract-signing protocol [22], the proposed protocol relies on long-term stable quantum memories, namely keeping entangled pairs until the binding phase. The practical problem of long-term stable quantum memories can be overcome by a simple modification: instead of EPR pairs, Trent sends (over an authenticated quantum channel) two ordered sets of photons in pure BB84 states, one for Alice and the other for Bob. As soon as the clients receive their particles, they measure on them the honest observables according to $h(M)$. Thus, all the information kept by the parties is classical and can be used to check the agents' behavior and honesty, similarly as in the protocol we proposed. This approach goes along the lines of reducing the security of BB84 to the E91 entangled QKD protocol. The details of such a reduction are a matter of a separate study.

Contract signing is an important and wide-spread cryptographic protocol, as performing transactions over the Internet is an important part of today's society. Nevertheless, all such current classical implementations fully rely on the use of a third trusted party. Moreover, if buying directly from, say, Amazon, the trust is handed over to a signing party, Amazon itself. Within classical approaches, there are several service providers and applications that can mediate the contract signing process, such as EverSign, HelloSign, and DocuSign.

On the other hand, if the parties do not want to rely fully on a third party to exchange the signature between them, classical solutions, such as [15–21], mentioned in the Introduction, require the gradual exchange of signed parts of the message. Therefore, classical digital contract signing is a very demanding application from the point of view of communication and computation: the exchange process has a significant number of rounds, and each exchanged message has to be digitally signed using (computationally-demanding) public-key cryptographic systems and infrastructures. We refer the reader to the following survey [27] for a more detailed description of these protocols. As a consequence, such classical contract signing applications are, due to their current inefficiency, to the best knowledge of the authors, not present on the market. We note that this is precisely the reason classical solutions to secure multiparty computation privacy protocols are not widely available as well. Potentially fast exchange of (single or coherent) photon pulses might be one of the main advantages of quantum solutions to the mentioned cryptographic problems. Note that by now, it was the higher

security levels that promoted quantum over classical cryptography. This novel feature, providing practical schemes not even available classically, might potentially be shown to be significant.

5. Conclusions

We presented a quantum protocol for signing contracts. We showed that, under the realistic assumptions of noise and measurement errors, the protocol was fair, and consequently optimistic as well. In particular, the maximal probability to cheat can be made arbitrarily low, as it scaled as $1/\sqrt{N}$, where $4N$ was the total number of rounds of the protocol. We also showed that our protocol was robust against noise. Indeed, even in a dishonest noisy case, the probability to cheat could be as low as 0.022, for $4N = 6000$, the noise parameter $\kappa = 0.05$, and $\alpha \in [0.7, 0.99]$. Moreover, already for $4N = 400$, the probability to cheat was below 10% (in the honest scenarios, this was so for $4N = 300$). Given the noise level κ, finding optimal parameters N, the range of α and $p(\alpha)$, was straightforward using the "brute force" techniques presented in the current paper (running computer codes and simulations). Full analytical study of such optimization problem is a matter of future study.

Unlike the classical counterparts, the protocol was based on the laws of physics, and the clients did not need to exchange a huge number of signed authenticated messages during the actual contract signing process (the exchange phase). Thus, the protocol was abuse-free: a client cannot prove to be involved in the actual act of signing the contract. In contrast, in the classical counterparts, having the signed and authenticated messages received during the exchange phase, a client could show them to other interested clients to negotiate better terms of a financial transaction. In other words, classically, a dishonest client can abuse the signing process by falsely presenting his/her interest in the deal, while actually using the protocol to achieve a different goal(s). Unlike generic quantum security protocols (say, quantum key distribution), preforming quantum measurements different from those prescribed by the protocol cannot help a cheating client (see Appendix A.2 for details). In the current proposal, each client can independently obtain the signed contract, without the other client being present, which was not possible in [22]. Thus, the probability to cheat was assigned to a real event (whereas in [22], it was just a formal figure of merit). Unlike the classical counterparts (and the previous quantum proposal [22]), when first contacting Trent, the clients did not need to agree upon a definitive contract. Moreover, Trent never learned the actual content of the protocol, as the clients provided Trent its hash value, $h^* = h(M)$, given by publicly-known hash function h.

Author Contributions: Conceptualization, N.P., P.Y., P.M., and A.S.; methodology, N.P., P.M., A.S., and P.Y.; formal analysis, P.Y., N.P., A.S., and P.M.; writing, original draft preparation, P.Y., N.P., and A.S.

Funding: The authors acknowledge the support of SQIG (Security and Quantum Information Group), the Instituto de Telecomunicações (IT) Research Unit, Ref. UID/EEA/50008/2019, funded by Fundação para a Ciência e Tecnologia (FCT), and the FCT projects Confident PTDC/EEI-CTP/4503/2014, QuantumMining POCI-01-0145-FEDER-031826, and Predict PTDC/CCI-CIF/29877/2017, supported by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme (COMPETE 2020), and by the Regional Operational Program of Lisbon. A.S. acknowledges funds granted to Laboratório de Sistemas Informáticos de Grande Escala (LASIGE) Research Unit, Ref. UID/CEC/00408/2013. P.Y. acknowledges the support of Doctoral Programme in the Physics and Mathematics of Information: Foundations of Future Information Technologies and FCT (Portugal) through the scholarship PD/BD/113648/2015.

Acknowledgments: The authors acknowledge Marie-Christine Röhsner for pointing out an attack on the protocol, which led to improvement of the work. N.P. acknowledges Aleksandra Dimić for fruitful discussions.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

Appendix A

In this Appendix, we analyze the protocol's security in case the communication is interrupted at step m. To proceed to the binding phase, the two clients measure their respective honest observables ($\hat{H}_{\mathtt{A}_i}$ and $\hat{H}_{\mathtt{B}_i}$) on the rest of their qubits, thus increasing their probability to pass Trent's Test 10a on

qubits shared with him. For the qubits shared between Trent and the other client, used in Test 10b, the best they can do is to guess their outcomes. Note that this strategy is optimal for both the case of honest clients, studied in Appendix A.1, as well as for a dishonest client, discussed in Appendix A.2.

Before discussing the protocol's correctness for the case of honest and security for dishonest clients, we present the overall setup: the "configuration" $\mathcal{L}$ and the details of the structure of the figure of merit, the expected probability to cheat for Bob and Alice. We assume the scenario in which the communication is interrupted upon both clients receiving m measurement outcomes from the other. The biased case of one client having m, while the other $m-1$, results is described analogously. A simplified protocol description is given in Figure A1.

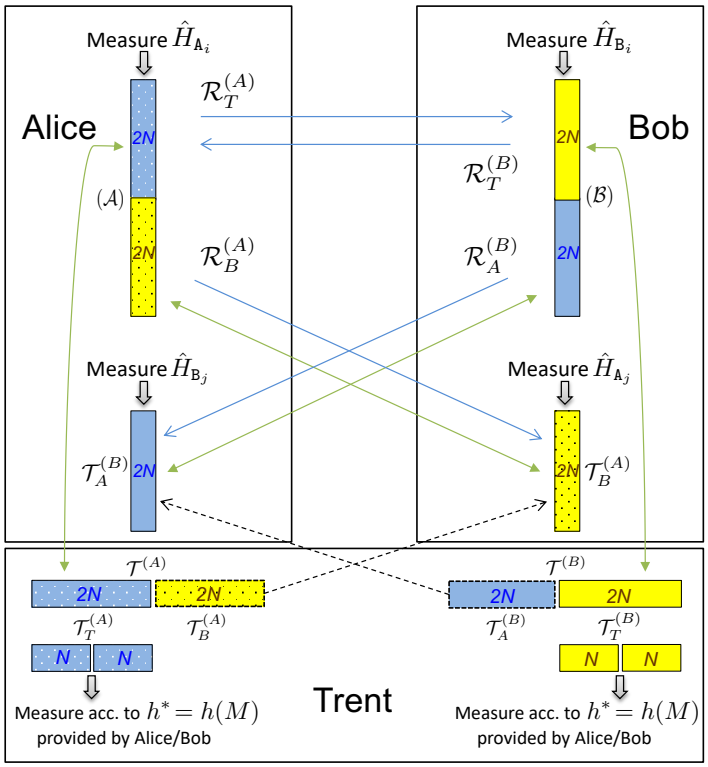


Figure A1. Stages of the protocol. The green arrows represent entanglement between the corresponding qubits. The blue arrows represent the transfer of measurement outcomes. The dashed arrows represent the transfer of qubits $\mathcal{T}_A^{(B)}$ and $\mathcal{T}_B^{(A)}$ from Trent to Alice and Bob, respectively. The big boxes represent the shielded private laboratories of Alice, Bob, and Trent.

Appendix A.1. Probabilistic Fairness for Honest Clients

In this scenario, we assume that both Alice and Bob are honest until step m, after which the communication is interrupted. Without loss of generality, we assume that Alice was the first one to start the information exchange. To analyze the fairness of the protocol, we consider the case when the communication is interrupted after Alice has sent m outcomes to Bob, but received only $m-1$ outcomes in return, so Bob has a slight advantage over Alice. The case where both Alice and Bob have m outcomes each is perfectly symmetric (that is, both Bob and Alice have the same expected probability to bind). In this section, we will analyze the ideal case of a noiseless channel, as well as the protocol's robustness in the presence of noise.

Appendix A.1.1. Noiseless Channel

In the absence of any noise, both honest Alice and Bob obtain perfect correlations on their respective qubits shared (entangled) with Trent and among themselves. Hence, both clients will, with certainty, pass each others' verification tests, $P_{ABS}(m) = P_{BAS}(m) = 1$, and thus, $P(m) =$

$P_{ABS}(m) \cdot P_{BAS}(m) = 1$ (i.e., the reason for communication interruption is the failure of the network). Then, the probability to cheat, say, for Bob (10), is given by:

$$P^{\mathcal{B}}_{ch}(m;\alpha) = P^{\mathcal{B}}_{bind}(m;\alpha)\left[1 - P^{\mathcal{A}}_{bind}(m;\alpha)\right]. \tag{A1}$$

Moreover, the probabilities for the clients to pass Trent's test on their own qubits, $P_{ATH}(m;\alpha)$ and $P_{BTH}(m;\alpha)$, are also one. Thus, Bob's probability to bind (9) is also simplified, giving:

$$P^{\mathcal{B}}_{bind}(m;\alpha) = P_{BTH}(m;\alpha)P_{BTA}(m;\alpha) = P_{BTA}(m;\alpha) \tag{A2}$$

and analogously for Alice.

Out of the m correct results that Bob received from Alice (note that we assume the ideal noiseless scenario), $N - \ell^{(A)}_{T2}$ are relevant when presenting to Trent to prove Alice's commitment. With α chosen randomly by Trent, let $\lfloor \alpha N \rfloor$ be the number of correct results corresponding to N qubits from $\mathcal{T}^{(A)}_{T2}$ and $\mathcal{T}^{(B)}_{T2}$ each that Trent needs to receive from Bob in order to bind the contract for him. For the case $\lfloor \alpha N \rfloor \leq N - \ell^{(A)}_{T2}$, Bob already has more than the required number of correct results; hence, his probability to bind the contract will be one. When $\lfloor \alpha N \rfloor > N - \ell^{(A)}_{T2}$, Bob must correctly guess at least $n_c = \lfloor \alpha N \rfloor - \left(N - \ell^{(A)}_{T2}\right)$ out of the remaining $\ell^{(A)}_{T2}$ results to convince Trent to bind the contract for him. Hence, his probability to bind the contract is given by:

$$P^{\mathcal{B}}_{bind}(m;\alpha|\ell^{(A)}_{T2}) = P_{BTA}(m;\alpha|\ell^{(A)}_{T2}) = \begin{cases} 1 & \text{if } \lfloor \alpha N \rfloor \leq N - \ell^{(A)}_{T2} \\ 2^{-\ell^{(A)}_{T2}} \sum\limits_{n=n_c}^{\ell^{(A)}_{T2}} \binom{\ell^{(A)}_{T2}}{n} & \text{if } \lfloor \alpha N \rfloor > N - \ell^{(A)}_{T2}. \end{cases} \tag{A3}$$

Here, $\binom{\ell^{(A)}_{T2}}{n}$ gives the number of possible choices for n out of $\ell^{(A)}_{T2}$, and $2^{-\ell^{(A)}_{T2}} = 2^{-n}\, 2^{-(\ell^{(A)}_{T2} - n)}$ gives the probability of guessing correctly exactly n results (incorrectly on the rest of the $\ell^{(A)}_{T2} - n$ results).

Averaging over $\ell^{(A)}_{T2}$ gives the binding probability for Bob, as a function of measurement outcomes obtained from Alice, m (the round of communication interruption), and the parameter α:

$$P^{\mathcal{B}}_{bind}(m;\alpha) = \sum_{\ell^{(A)}_{T2}} q\left(\ell^{(A)}_{T2}\right) P^{\mathcal{B}}_{bind}\left(m;\alpha|\ell^{(A)}_{T2}\right), \tag{A4}$$

with the probability distribution for $\ell^{(A)}_{T2}$ given by:

$$q\left(\ell^{(A)}_{T2}\right) = \frac{\binom{m}{N - \ell^{(A)}_{T2}}\binom{4N - m}{\ell^{(A)}_{T2}}}{\binom{4N}{N}}. \tag{A5}$$

To verify the last expression, note that at the m-th step: (i) there are $\binom{m}{N - \ell^{(A)}_{T2}}$ ways of choosing the $N - \ell^{(A)}_{T2}$ results already obtained from Alice; (ii) $\binom{4N - m}{\ell^{(A)}_{T2}}$ ways of choosing the $\ell^{(A)}_{T2}$ results from the remaining $4N - m$ to guess; (iii) there are $\binom{4N}{N}$ ways of choosing the N results, relevant for binding the contract for Bob, from the total $4N$ results.

Note that in general, $\ell_{T2}^{(A)}$ takes values from 0–N, but for the step m of communication interruption, $\ell_{T2}^{(A)}$ is constrained to the following values:

$$\begin{aligned} \ell_{T2}^{(A)} &= N-m\,,\ldots,N && \text{for } m<N, \\ \ell_{T2}^{(A)} &= 0\,,\ldots,N && \text{for } N\leq m<3N, \\ \ell_{T2}^{(A)} &= 0\,,\ldots,4N-m && \text{for } 3N\leq m. \end{aligned} \tag{A6}$$

The above ranges define our summation $\sum_{\ell_{T2}^{(A)}}$. When $\ell_{T2}^{(A)}$ is out of the first or the third range, we have $m<N-\ell_{T2}^{(A)}$ and $4N-m<\ell_{T2}^{(A)}$, respectively. Since in those cases, the binomials $\binom{m}{N-\ell_{T2}^{(A)}}$ and $\binom{4N-m}{\ell_{T2}^{(A)}}$ are by definition equal to zero, we can always take the $\ell_{T2}^{(A)}=0\,,\ldots,N$ range.

By replacing m with $m-1$ in Equation (A5), one obtains Alice's probability to bind the contract $P_{bind}^{\mathcal{A}}(m;\alpha)$ (note that, having one less measurement result than Bob, she therefore has a small disadvantage).

Bob's probability to cheat (A1), when the communication is interrupted at step m, and for a fixed α, is given by $P_{ch}^{\mathcal{B}}(m;\alpha)=P_{bind}^{\mathcal{B}}(m;\alpha)\left[1-P_{bind}^{\mathcal{A}}(m;\alpha)\right]$. Using (A4), and its counterpart for Alice, one can evaluate the expected probability to cheat $\bar{P}_{ch}^{\mathcal{B}}(m)=\int p(\alpha)\,P_{ch}^{\mathcal{B}}(m;\alpha)\,\mathrm{d}\alpha$ for every m. In Figure A2, we plot the expected probability to cheat, $\bar{P}_{ch}^{\mathcal{B}}(m)$, against the step of communication interruption m (running from 1–4N), for the simplest case of the uniform distribution $p(\alpha)$ on the interval $[0.9,0.99]$. Since Alice starts first and Bob is thus privileged, we have that $\bar{P}_{ch}^{\mathcal{A}}(m)\leq\bar{P}_{ch}^{\mathcal{B}}(m)=\bar{P}_{ch}(m)$. The maximal value shows the behavior $\max_m\bar{P}_{ch}(m)\propto N^{-1/2}$, as presented in Figure 6a,d, from the main text.

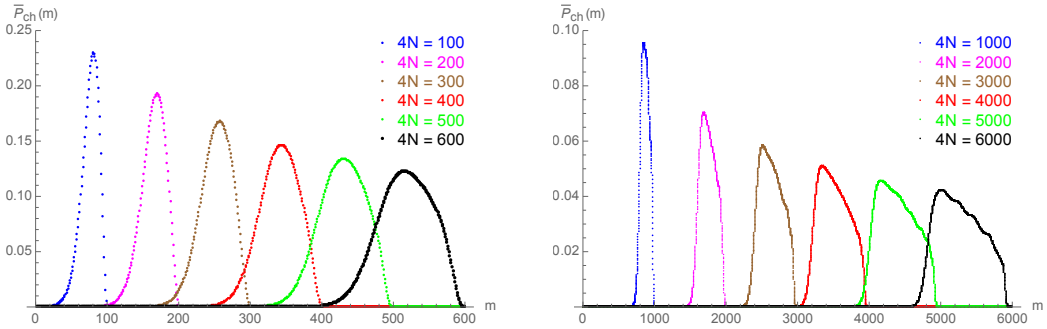


Figure A2. Honest scenario (noiseless channel): the expected probability to cheat, $\bar{P}_{ch}(m)$, is plotted against the step of communication interruption m (from 1 to 4N), for α chosen uniformly on the interval $[0.9,0.99]$.

In addition to the quantitative results for up to $4N=6000$, below, we present an analytic proof of the asymptotic behavior for the maximal expected probability to cheat, showing that as $N\to\infty$, we have $P_{ch}(m;\alpha)\propto N^{-1/2}$. The proof for the case of a noisy channel follows analogously (see [28] for the analysis for the contract signing presented in [22]).

Theorem A1. *In the honest noiseless case, for uniformly-chosen $\alpha\in(1/2,1)$, we have that:*

$$P_{ch}(m;\alpha)\in\mathcal{O}(N^{-1/2}).$$

Proof. We are going to show that for all $1<m<4N$ and $\alpha\in(1/2,1)$, we have that $P_{ch}(m;\alpha)\in\mathcal{O}(N^{-1/2})$, and therefore, the statement follows straightforwardly.

Recall that a binomial distribution $B(k,p)$, with k sufficiently large and p bounded away from zero and one (that is, not tending to zero or one as N grows to infinity), can be approximated by a normal

distribution with mean, kp, and variance, $kp(1-p)$, as $\mathcal{N}(kp, kp(1-p))$. In our case (with $k = \ell_{T2}^{(A)}$), for $\ell_{T2}^{(A)}$ sufficiently large and $p = 1/2$ (guessing probability of a binary bit), the binomial distribution, $B\left(\ell_{T2}^{(A)}, 1/2\right)$ can be approximated by $\mathcal{N}\left(\ell_{T2}^{(A)}/2\,,\ \ell_{T2}^{(A)}/4\right)$. Hence, approximating $P_{bind}^{\mathcal{B}}\left(m;\alpha|\ell_{T2}^{(A)}\right)$ for the case $\lfloor \alpha N \rfloor > N - \ell_{T2}^{(A)}$, we have:

$$\begin{aligned} P_{bind}^{\mathcal{B}}(m;\alpha|\ell_{T2}^{(A)}) &= 2^{-\ell_{T2}^{(A)}} \sum_{n=n_c}^{\ell_{T2}^{(A)}} \binom{\ell_{T2}^{(A)}}{n} \\ &\approx \int_{n_c}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(n-\mu_B)^2}{2\sigma_B^2}} \mathrm{d}n = \frac{1}{2}\mathrm{erfc}\left[\sqrt{\frac{2}{\ell_{T2}^{(A)}}}\left(n_c - \frac{\ell_{T2}^{(A)}}{2}\right)\right], \end{aligned} \tag{A7}$$

with $n_c = \lfloor \alpha N \rfloor - (N - \ell_{T2}^{(A)})$ and the binomial mean and variance given by $\mu_B = \ell_{T2}^{(A)}/2$ and $\sigma_B^2 = \ell_{T2}^{(A)}/4$, respectively. Therefore:

$$P_{bind}^{\mathcal{B}}(m;\alpha|\ell_{T2}^{(A)}) \approx \begin{cases} 1 & \text{if } \lfloor \alpha N \rfloor \leq N - \ell_{T2}^{(A)}\,, \\ \frac{1}{2}\ \mathrm{erfc}\left[\sqrt{\frac{2}{\ell_{T2}^{(A)}}}\left(N(\alpha-1) + \frac{\ell_{T2}^{(A)}}{2}\right)\right] & \text{if } \lfloor \alpha N \rfloor > N - \ell_{T2}^{(A)}\,, \end{cases} \tag{A8}$$

where $\mathrm{erfc}(\beta)$ is the complementary error function, defined as:

$$\mathrm{erfc}(\beta) = \frac{2}{\sqrt{\pi}} \int_{\beta}^{\infty} e^{-\gamma^2} \mathrm{d}\gamma\,. \tag{A9}$$

For the case $\lfloor \alpha N \rfloor > N - \ell_{T2}^{(A)}$, we expand $P_{bind}^{\mathcal{B}}(m;\alpha|\ell_{T2}^{(A)})$ around $\ell_{T2}^{(A)} = 2N(1-\alpha)$ where it has the value $1/2$, to obtain:

$$\begin{aligned} \frac{1}{2}\ \mathrm{erfc}\left[\sqrt{\frac{2}{\ell_{T2}^{(A)}}}\left(N(\alpha-1) + \frac{\ell_{T2}^{(A)}}{2}\right)\right] &= \frac{1}{2} + c_1(\ell_{T2}^{(A)} - 2N(1-\alpha)) + c_2(\ell_{T2}^{(A)} - 2N(1-\alpha))^2 \\ &\quad + \mathcal{O}((\ell_{T2}^{(A)} - 2N(1-\alpha))^3). \end{aligned} \tag{A10}$$

Let $P_1(m;\alpha|\ell_{T2}^{(A)}) = \frac{1}{2} + c_1(\ell_{T2}^{(A)} - 2N(1-\alpha)) + c_2(\ell_{T2}^{(A)} - 2N(1-\alpha))^2$. We also obtain the limits for $\ell_{T2}^{(A)}$ where this function becomes one and zero, denoted by ℓ_{inf} and ℓ_{sup}, respectively.

We know that:

$$P_{bind}^{\mathcal{B}}(m;\alpha) = \sum_{\ell_{T2}^{(A)}} q\left(\ell_{T2}^{(A)}\right) P_{bind}^{\mathcal{B}}\left(m;\alpha|\ell_{T2}^{(A)}\right), \tag{A11}$$

with the probability distribution for $\ell_{T2}^{(A)}$ given by:

$$q\left(\ell_{T2}^{(A)}\right) = \frac{\binom{m}{N-\ell_{T2}^{(A)}}\binom{4N-m}{\ell_{T2}^{(A)}}}{\binom{4N}{N}}. \tag{A12}$$

By Feller's result [29] on the approximation of the hypergeometric distribution, one can approximate (A12) for $N \to \infty$, $\frac{N}{4N} \to t \in (0,1)$ to a normal distribution as:

$$q(\ell_{T2}^{(A)}) \sim \frac{e^{-(\ell_{T2}^{(A)} - \mu_\ell)^2/2\sigma_\ell^2(1-t)}}{\sqrt{2\pi\sigma_\ell^2(1-t)}} = P_2(\ell_{T2}^{(A)})\,, \tag{A13}$$

with mean, $\mu_\ell = N - m/4$, variance, $\sigma_\ell^2 = N(\frac{m}{4N})(\frac{4N-m}{4N})$, and $t = 1/4$. Hence, (A11) can be upper bounded (shifting the limit $\ell_{T2}^{(A)} = N(1-\alpha)$ to $\ell_{T2}^{(A)} = \ell_{inf}$ and using the approximation (A10) from ℓ_{inf} to ℓ_{sup}) by:

$$P_{bind}^{\mathcal{B}}(m;\alpha) \leq \int_0^{\ell_{inf}} 1 \cdot P_2(\ell_{T2}^{(A)})\, \mathrm{d}\ell_{T2}^{(A)} + \underbrace{\int_{\ell_{inf}}^{\ell_{sup}} P_2(\ell_{T2}^{(A)}) \cdot \left[P_1(m;\alpha|\ell_{T2}^{(A)}) + k\left(\ell_{T2}^{(A)} - 2N(1-\alpha)\right)^3 \right] \mathrm{d}\ell_{T2}^{(A)}}_{B_1}$$
$$+ \underbrace{\int_{\ell_{sup}}^{+\infty} P_2(\ell_{T2}^{(A)}) \cdot P_{bind}^{\mathcal{B}}(m;\alpha|\ell_{T2}^{(A)})\, \mathrm{d}\ell_{T2}^{(A)}}_{B_2} \tag{A14}$$

$$\leq \int_0^{\ell_{inf}} P_2(\ell_{T2}^{(A)})\, \mathrm{d}\ell_{T2}^{(A)} + \mathcal{O}(N^{-1/2}), \tag{A15}$$

where the $\mathcal{O}(N^{-1/2})$ in Equation (A15) was obtained by computing limits for B_1 and B_2 for $N \to \infty$ with the Mathematica software package. To compute the limit for B_1, from [30], we used the fact that by the change of variable $N = 1/x$, if $\lim_{x\to 0^+} f(x) = 0$, then $\lim_{N\to\infty} f(N) = 0$ and obtained the desired bound. For computing B_2, we upper bounded $P_{bind}^{\mathcal{B}}(m;\alpha|\ell_{T2}^{(A)})$ by $\frac{1}{2}\mathrm{erfc}[\lambda N]$ for some constant $\lambda > 0$ and used the upper bound for complementary error function, which also led to the $O(N^{-1/2})$ bound.

To obtain the expression for Alice's probability to bind, $P_{bind}^{\mathcal{A}}(m;\alpha)$, we have:

$$q\left(\ell_{T1}^{(B)}\right) = \frac{\binom{m-1}{N-\ell_{T1}^{(B)}}\binom{4N-m+1}{\ell_{T1}^{(B)}}}{\binom{4N}{N}}. \tag{A16}$$

Making the same approximations, we get $P_2\left(\ell_{T1}^{(B)}\right)$ similar to (A13), with m replaced by $m-1$. Hence:

$$P_{bind}^{\mathcal{A}}(m;\alpha) \geq \int_0^{\ell_{inf}} P_2(\ell_{T1}^{(B)})\, \mathrm{d}\ell_{T1}^{(B)} - \underbrace{\int_0^{\ell_{inf}} (1 - P_{bind}^{\mathcal{A}}(m;\alpha|\ell_{T1}))P_2(\ell_{T1}^{(B)})}_{A_1}\, \mathrm{d}\ell_{T1}^{(B)}$$
$$+ \underbrace{\int_{\ell_{inf}}^{+\infty} P_2(\ell_{T1}^{(B)})\, P_{bind}^{\mathcal{A}}(m;\alpha|\ell_{T1}^{(B)})\, \mathrm{d}\ell_{T1}^{(B)}}_{A_2} \tag{A17}$$

$$\geq \int_0^{\ell_{inf}} P_2(\ell_{T1}^{(B)})\, \mathrm{d}\ell_{T1}^{(B)} - \mathcal{O}(N^{-1/2}), \tag{A18}$$

where the bounds in Equation (A18) were obtained with the help of the Mathematica software package. Since we are computing the lower bound, one can drop A_2, and it suffices to show that the integral of A_1 computed between $N(1-\alpha)$ (as $P_{bind}^{\mathcal{A}}(m;\alpha|\ell_{T1})$ is one up to $N(1-\alpha)$) and ℓ_{inf} vanishes when $N \to \infty$. Since $(1 - P_{bind}^{\mathcal{A}}(m;\alpha|\ell_{T1})) < 1$, we can drop this and compute the limit of the integral of $P_2(\ell_{T1}^{(B)})$, which also provides the bound.

If the communication was interrupted at step m, for all choices of $\alpha \in (1/2, 1)$ by Trent, we want to compute, say for Bob, the probability to cheat, given by:

$$P_{ch}^{\mathcal{B}}(m;\alpha) = P_{bind}^{\mathcal{B}}(m;\alpha)\left[1 - P_{bind}^{\mathcal{A}}(m;\alpha)\right] \tag{A19}$$
$$\leq \left[\int_0^{\ell_{inf}} P_2(\ell_{T2}^{(A)})\, \mathrm{d}\ell_{T2}^{(A)} + O(N^{-\frac{1}{2}})\right]\left[1 - \int_0^{\ell_{inf}} P_2(\ell_{T1}^{(B)})\, \mathrm{d}\ell_{T1}^{(B)} + \mathcal{O}(N^{-1/2})\right] \tag{A20}$$
$$= \mathcal{O}(N^{-1/2}), \tag{A21}$$

where the $\mathcal{O}(N^{-1/2})$ in Equation (A21) was again obtained by computing a limit with the Mathematica software package. □

Appendix A.1.2. Noisy Channel

For the case of a noisy channel, a binomial test could be used by both parties for the permitted number of wrong results from the other party. Consider white noise in the channel that decreases the degree of correlation between the honest clients' results. The depolarizing channel, modeling the effects of white noise on a two-qubit mixed state ρ, is given by (I is the identity matrix for dimension four, and $\kappa \in [0,1]$):

$$\mathcal{E}_d(\rho) = (1-\kappa)\rho + \kappa\frac{I}{4}. \tag{A22}$$

Hence, both Alice and Bob receive some inevitable number of incorrect results, due to the noise considered, in spite of both of them measuring their respective honest observables on their own qubits. Let us denote by $p(xy)$ the probability of Alice and Bob obtaining the results x and y, respectively, corresponding to measurements of $\mathcal{E}_d(\rho)$ in either a computational or diagonal basis, for the case of the entangled two-qubit state $\rho = |\psi^+\rangle\langle\psi^+|$. Given that both Alice and Bob measure their respective correct observables, $\hat{H}_{\mathtt{A}_i}$ and $\hat{H}_{\mathtt{B}_i}$, we have the following probabilities of obtaining different results:

$$\begin{aligned} p(00) &= (2-\kappa)/4, \\ p(01) &= \kappa/4, \\ p(10) &= \kappa/4, \\ p(11) &= (2-\kappa)/4. \end{aligned} \tag{A23}$$

For a given probability of the favorable event (in our case, $p(00)+p(11)$, of obtaining consistent results) and the total number of such events (say, m_r), the probability of obtaining exactly r correct results is given by the binomial distribution:

$$\mathcal{P}\left(r||\psi^+\rangle; m_r, \kappa\right) = \binom{m_r}{r}\Big(p(00)+p(11)\Big)^r\Big(p(01)+p(10)\Big)^{m_r-r}, \tag{A24}$$

with mean and variance given by:

$$\begin{aligned} \mu &= m_r\Big(p(00)+p(11)\Big) = m_r\left(1-\frac{\kappa}{2}\right), \\ \sigma^2 &= m_r\Big(p(00)+p(11)\Big)\Big(1-p(00)-p(11)\Big) = m_r\left(1-\frac{\kappa}{2}\right)\frac{\kappa}{2}. \end{aligned} \tag{A25}$$

Suppose that each party applies a three-sigma "acceptance criterion", then a client will continue as long as he/she has at least $\mu - 3\sigma$ consistent results from the other.

From (9), Bob's probability to bind the contract is $P^{\mathcal{B}}_{bind}(m;\alpha) = \langle P_{BTH}(m;\alpha|\ell^{(B)}_{T2})\rangle_{\ell^{(B)}_{T2}} \cdot \langle P_{BTA}(m;\alpha|\ell^{(A)}_{T2})\rangle_{\ell^{(A)}_{T2}}$. While Bob's probability to pass the test on his qubits, P_{BTH}, does not depend on step m nor on $\ell^{(B)}_{T2}$ (he measures all of his qubits), the probability to pass the test on Alice's qubits, P_{BTA}, depends on m, as well as on $\ell^{(A)}_{T2}$ (for simplicity, we omit the implicit dependence on the noise parameter κ). Thus, we have:

$$P^{\mathcal{B}}_{bind}(m;\alpha) = P_{BTH}(\alpha)\cdot\langle P_{BTA}(m;\alpha|\ell^{(A)}_{T2})\rangle_{\ell^{(A)}_{T2}}, \tag{A26}$$

with:

$$P_{BTH}(\alpha) = \underbrace{\sum_{s=\lfloor \alpha N \rfloor}^{N} \binom{N}{s} P_{=}^{s} P_{\neq}^{N-s}}_{\substack{\text{probability of obtaining at least} \\ \lfloor \alpha N \rfloor \text{ correct results from } N \text{ by measuring} \\ \text{the correct observable } (\hat{H}_{B_i}) \text{ on all his qubits}}}, \tag{A27}$$

$$P_{BTA}(m;\alpha) = \sum_{\ell_{T2}^{(A)}} q\left(\ell_{T2}^{(A)}\right) \underbrace{\sum_{t=0}^{N-\ell_{T2}^{(A)}} \binom{N-\ell_{T2}^{(A)}}{t} P_{=}^{t} P_{\neq}^{(N-\ell_{T2}^{(A)})-t}}_{\substack{\text{probability of obtaining } t \text{ correct} \\ \text{results from the } N-\ell_{T2}^{(A)} \text{ results} \\ \text{received up to step } m}} \times \underbrace{2^{-\ell_{T2}^{(A)}} \sum_{u=\lfloor \alpha N \rfloor - t}^{\ell_{T2}^{(A)}} \binom{\ell_{T2}^{(A)}}{u}}_{\substack{\text{probability of guessing} \\ \text{at least } \lfloor \alpha N \rfloor - t \text{ correct results} \\ \text{from the } \ell_{T2}^{(A)} \text{ results to guess}}}, \tag{A28}$$

where $P_{=} = p(00) + p(11)$ and $P_{\neq} = p(10) + p(01)$. Analogously, we define $P_{bind}^{\mathcal{A}}(m;\alpha) = P_{ATH}(\alpha) \cdot \langle P_{ATB}(m;\alpha|\ell_{T1}^{(B)})\rangle_{\ell_{T1}^{(B)}}$ for Alice.

To obtain Bob's probability to cheat (10), $P_{ch}^{\mathcal{B}}(m;\alpha) = P(m) P_{bind}^{\mathcal{B}}(m;\alpha)\left[1 - P_{bind}^{\mathcal{A}}(m;\alpha)\right]$, we now estimate $P(m) = P_{ABS}(m) \cdot P_{BAS}(m)$. $P_{BAS}(m)$, the probability that Bob passes Alice's tests on their shared qubits (from the $m-1$ results sent by him to Alice). Note that, in order to reach step m, Bob has to pass the test at all the steps $2, \ldots, m-2, m-1$. Thus, we can bound $P_{BAS}(m)$ from above by the probability to pass Alice's test at step $m-1$ only:

$$\mathcal{P}_{BAS}(m) = \sum_{\ell_A^{(B)}} q(\ell_A^{(B)}) \times \underbrace{\sum_{v=\mu-3\sigma}^{2N-\ell_A^{(B)}} \binom{2N-\ell_A^{(B)}}{v} P_{=}^{v} P_{\neq}^{2N-\ell_A^{(B)}-v}}_{\substack{\text{probability for Alice to obtain at least} \\ \mu-3\sigma \text{ correct results on the } 2N-\ell_A^{(B)} \text{ results} \\ \text{out of the } m-1 \text{ results received from Bob}}}, \tag{A29}$$

with μ and σ defined by (A25), where $m_r = 2N - \ell_A^{(B)}$. The probability distribution for $\ell_A^{(B)}$ is given by:

$$q\left(\ell_A^{(B)}\right) = \frac{\binom{m-1}{2N-\ell_A^{(B)}}\binom{4N-m+1}{\ell_A^{(B)}}}{\binom{4N}{2N}}. \tag{A30}$$

$P_{ABS}(m)$, the probability that Alice passes Bob's test on their shared qubits (from the m results she sends to Bob), as well as its bound $\mathcal{P}_{ABS}(m)$ are defined analogously. Thus, the overall probability to reach step m is given by $P_{ABS}(m) \cdot P_{BAS}(m) \leq \mathcal{P}_{ABS}(m) \cdot \mathcal{P}_{BAS}(m)$.

Hence, the average probability for Bob to cheat, with both Alice and Bob having reached the step m by passing each others' tests, is

$$\bar{P}_{ch}^{\mathcal{B}}(m) \leq \mathcal{P}_{ABS}(m) \cdot \mathcal{P}_{BAS}(m) \int_{\alpha} p(\alpha) P_{bind}^{\mathcal{B}}(m;\alpha)\left(1 - P_{bind}^{\mathcal{A}}(m;\alpha)\right) \, \mathrm{d}\alpha. \tag{A31}$$

In Figure A3, we present the expected probability to cheat, $\bar{P}_{ch}(m)$, plotted against the total number of communications between Alice and Bob, $4N$, for the three-sigma acceptance criterion of both Alice and Bob (as before, $p(\alpha)$ is uniform on the interval $[0.9, 0.99]$). The maximal value again shows the behavior $\max_m \bar{P}_{ch}(m) \propto N^{-1/2}$, as presented in Figure 6b,e from the main text.

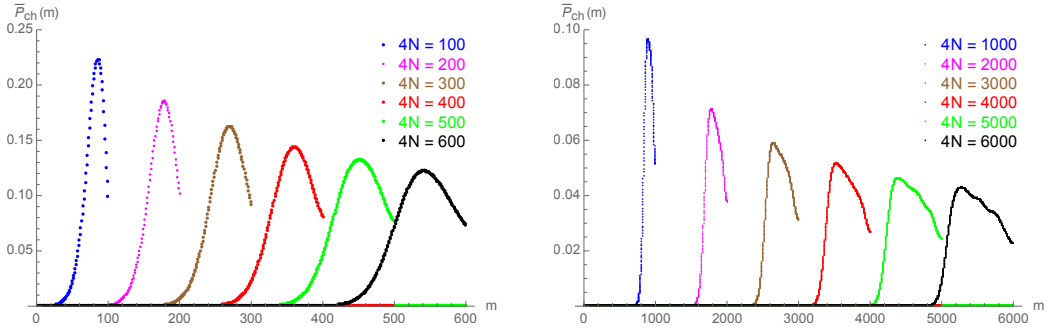


Figure A3. Honest scenario (noisy channel): the expected probability to cheat, $\bar{P}_{ch}(m)$, is plotted against the step of communication interruption m (from 1–$4N$), for α chosen uniformly over the interval $[0.9, 0.99]$. The noise parameter is $\kappa = 0.05$.

Appendix A.2. Security Analysis (Probabilistic Fairness) against a Dishonest Client

In order to cheat, a dishonest client, say Bob, would want to obtain a signed copy of the contract for message M that Alice and Bob initially agreed upon, without letting Alice obtain a signed copy for herself, so that he can use it later on, if he wants.

The qubits from $\mathcal{T}_B^{(A)}$ are used to verify Alice's honesty, and therefore, Bob measures them according to the protocol (note that he knows which qubits from $\mathcal{A}$ are entangled with $\mathcal{T}_B^{(A)}$ and also which observable $H_{\mathtt{A}_i}$ to measure). Regarding qubits from $\mathcal{B}$, unlike the standard quantum cryptographic protocols, where the task of an adversary (say Eve, in key distribution schemes) is to distinguish between mutually non-orthogonal states, in this protocol, a cheating Bob knows the pure states of his qubits. He knows that Trent and an honest Alice measure $\hat{H}_{\mathtt{B}_i}$ on their respective halves of the pairs entangled with Bob, and thus, by measuring $\hat{H}_{\mathtt{B}_i}$, he can check to which of the two mutually-orthogonal states they (will) collapse his system. Note that Bob cannot take advantage of measuring observables different from those prescribed by the protocol. Since Alice's and Bob's measurements are local, they commute, and thus, their time order is irrelevant. Thus, by deviating from the protocol, a dishonest Bob can only spoil the correlations between his and Alice's outcomes, a task he can achieve by any random source. Regarding Trent, the time ordering matters, as it is Bob who tells Trent which single-qubit observables to measure, based on the $h^* = h(M)$ he provides Trent. Nevertheless, since Bob's aim is to have his results as correlated as possible with Trent's, without spoiling the correlations established with Alice's results, Bob should measure his honest observables $\hat{H}_{\mathtt{B}_i}$ on all the qubits from $\mathcal{B}$.

As argued above, for a cheating client Bob, even knowing to which states the particles sent to him are collapsed to (due to Alice's and Trent's measurements) does not help. This is because he still does not know which of those particles are entangled with Alice and which ones are with Trent. Bob's cheating strategy should allow him to bind the contract, such that Alice is unable to bind the contract even with Trent's help. In order to bind the contract for himself, Bob must pass the test by providing to Trent at least a fraction α of correct measurement results on the qubits entangled with $\mathcal{T}_{T2}^{(A)}$, as well as $\mathcal{T}_{T2}^{(B)}$, respectively. At the same time, Bob does not want Alice to pass the test on the results sent by him (corresponding to qubits entangled with $\mathcal{T}_{T1}^{(B)}$) to her. Making measurements in a basis other than the one given by $\hat{H}_{\mathtt{B}_i}$ on some of the qubits and sending those results to Alice does not help him in any way, because then he ends up with incorrect results on some of his qubits in $\mathcal{T}_{T2}^{(B)}$ entangled with Trent. While these incorrect results will spoil Alice's chances to bind the contract, they will equally decrease Bob's chances as well. Hence, Bob's best strategy is to measure his honest observable $\hat{H}_{\mathtt{B}_i}$ on all his qubits, to pass the test on all $\mathcal{T}_{T2}^{(B)}$ qubits, and send random bits to Alice, by choosing a random

probability f to decide whether or not to flip the result that he sends to Alice (the frequency of sending wrong results). Therefore, the probability for Alice to obtain correct results on these qubits will be:

$$\tilde{P}_{=} = (1-f)P_{=} + fP_{\neq}, \tag{A32}$$

where $P_{=} = p(00) + p(11)$ and $P_{\neq} = p(10) + p(01)$.

The communication is interrupted at step m, after Alice stops sending the measurement outcomes to Bob upon suspecting a dishonest behavior. Recall that we are considering the case where Alice is the first one to start the communication, and therefore, in the worst case scenario after the m-th step, Alice and Bob each have m measurement results from each other.

With m measurement results each, the probability for Alice and Bob to pass Trent's test on her/his own qubits, $P_{ATH}(\alpha)$ and $P_{BTH}(\alpha)$, respectively, remains the same as in the honest noisy case. Since Bob decides to flip the results randomly, based on f, he is bound to send wrong results on some of the qubits. Hence, Alice receives more incorrect results from Bob as compared to the honest noisy case. Bob on the other hand receives the same number of correct results as in the honest noisy case. Therefore, $\mathcal{P}_{ABS}(m)$ and $P_{BTA}(m;\alpha)$ remain the same, while $P_{ATB}(m;\alpha)$ and $\mathcal{P}_{BAS}(m)$ are modified by replacing $P_{=}$ and $P_{\neq}$ by $\tilde{P}_{=}$ and $\tilde{P}_{\neq} = 1 - \tilde{P}_{=}$ (A32).

The expected probability to cheat, $\bar{P}_{ch}(m,f)$, is plotted in Figure A4. The same behavior (as in the case of honest clients), $\max_{m,f} \bar{P}_{ch}(m,f) \propto N^{-1/2}$, is observed, as in Figure 6c,f from the main text.

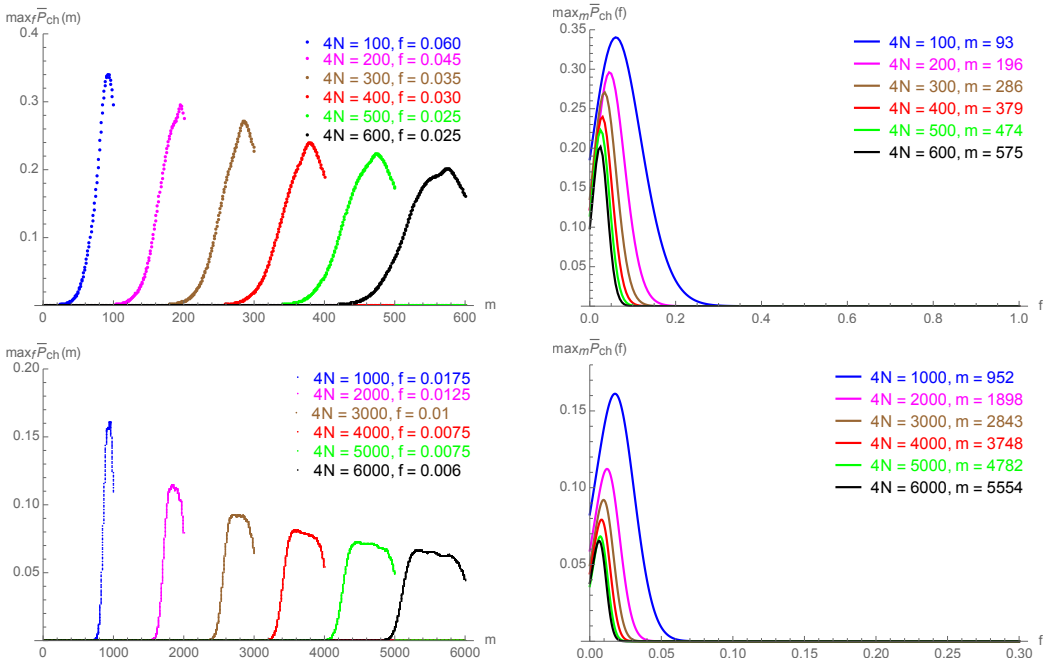


Figure A4. Dishonest scenario (noisy channel): The maximal expected probability to cheat, $\bar{P}_{ch}(m,f)$, is first plotted against the step of communication interruption, m (taking the optimal value for f), and then against the cheating parameter, f (taking the optimal value for m), for the case for the uniform $p(\alpha)$ on the interval $[0.9, 0.99]$. The noise parameter is $\kappa = 0.05$.

References

1. Wiesner, S. Conjugate coding. *ACM Sigact News* **1983**, *15*, 78–88. [CrossRef]
2. Bennett, C.H.; Brassard, G. Quantum cryptography: Public key distribution and coin tossing. *Theor. Comput. Sci.* **2014**, *560*, 7–11. [CrossRef]
3. Long, G.L.; Liu, X.S. Theoretically efficient high-capacity quantum-key-distribution scheme. *Phys. Rev. A* **2002**, *65*, 032302. [CrossRef]
4. Hillery, M.; Bužek, V.; Berthiaume, A. Quantum secret sharing. *Phys. Rev. A* **1999**, *59*, 1829. [CrossRef]

5. Giovannetti, V.; Lloyd, S.; Maccone, L. Quantum private queries. *Phys. Rev. Lett.* **2008**, *100*, 230502.
6. Fei, G.; Sujuan, Q.; Wei, H.; Wen, Q. Quantum private query: A new kind of practical quantum cryptographic protocols. *Sci. China Phys. Mech. Astron.* **2019**, *62*, 70301.
7. Sheng, Y.B.; Zhou, L. Distributed secure quantum machine learning. *Sci. Bull.* **2017**, *62*, 1025–1029. [CrossRef]
8. Lindell, Y.; Pinkas, B. Secure Multiparty Computation for Privacy-Preserving Data Mining. *J. Priv. Confid.* **2009**, 59–98. [CrossRef]
9. Broadbent, A.; Schaffner, C. Quantum cryptography beyond quantum key distribution. *Des. Codes Cryptogr.* **2016**, *78*, 351–382. [CrossRef]
10. Even, S.; Yacobi, Y. Relations among Public Key Signature Systems; Technical Report. 1980, p. 175, unpublished. Available online: https://www.researchgate.net/publication/243769259 (accessed on 10 July 2019).
11. Loura, R.; Almeida, A.J.; André, P.S.; Pinto, A.N.; Mateus, P.; Paunković, N. Noise and measurement errors in a practical two-state quantum bit commitment protocol. *Phys. Rev. A* **2014**, *89*, 052336. [CrossRef]
12. Almeida, A.J.; Stojanović, A.D.; Paunković, N.; Loura, R.; Muga, N.J.; Silva, N.A.; Mateus, P.; André, P.S.; Pinto, A.N. Implementation of a two-state quantum bit commitment protocol in optical fibers. *J. Opt.* **2016**, *18*, 015202. [CrossRef]
13. Loura, R.; Arsenović, D.; Paunković, N.; Popović, D.B.; Prvanović, S. Security of two-state and four-state practical quantum bit-commitment protocols. *Phys. Rev. A* **2016**, *94*, 062335. [CrossRef]
14. Fischer, M.J.; Lynch, N.A.; Paterson, M.S. Impossibility of distributed consensus with one faulty process. *J. ACM (JACM)* **1985**, *32*, 374–382. [CrossRef]
15. Even, S. A protocol for signing contracts. *ACM SIGACT News* **1983**, *15*, 34–39. [CrossRef]
16. Goldreich, O. A simple protocol for signing contracts. In *Advances in Cryptology*; Springer: Berlin/Heidelberg, Germany, 1984; pp. 133–136.
17. Even, S.; Goldreich, O.; Lempel, A. A randomized protocol for signing contracts. *Commun. ACM* **1985**, *28*, 637–647. [CrossRef]
18. Asokan, N.; Schunter, M.; Waidner, M. Optimistic protocols for fair exchange. In Proceedings of the 4th ACM Conference on Computer and Communications Security, Zurich, Switzerland, 1–4 April 1997; pp. 7–17.
19. Asokan, N.; Shoup, V.; Waidner, M. Optimistic fair exchange of digital signatures. In *Advances in Cryptology EUROCRYPT '98*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 591–606.
20. Ben-Or, M.; Goldreich, O.; Micali, S.; Rivest, R.L. A fair protocol for signing contracts. *IEEE Trans. Inf. Theory* **1990**, *36*, 40–46. [CrossRef]
21. Rabin, M.O. Transaction protection by beacons. *J. Comput. Syst. Sci.* **1983**, *27*, 256–267. [CrossRef]
22. Paunković, N.; Bouda, J.; Mateus, P. Fair and optimistic quantum contract signing. *Phys. Rev. A* **2011**, *84*, 062331. [CrossRef]
23. Situ, H.; Qiu, D.; Mateus, P.; Paunković, N. Secure N-dimensional simultaneous dense coding and applications. *Int. J. Quantum Inf.* **2015**, *13*, 1550051. [CrossRef]
24. Garay, J.A.; Jakobsson, M.; MacKenzie, P. Abuse-free optimistic contract signing. In Proceedings of the Annual International Cryptology Conference, Santa Barbara, CA, USA, 15–19 August 1999; pp. 449–466.
25. Stevens, M.; Bursztein, E.; Karpman, P.; Albertini, A.; Markov, Y. The First Collision for Full SHA-1. In *Advances in Cryptology—CRYPTO 2017*; Katz, J., Shacham, H., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2017; Volume 10401, pp. 570–596.
26. Boneh, D.; Dagdelen, Ö.; Fischlin, M.; Lehmann, A.; Schaffner, C.; Zhandry, M. Random Oracles in a Quantum World. In *Advances in Cryptology—ASIACRYPT 2011*; Lee, D.H., Wang, X., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2011; Volume 7073, pp. 41–69.
27. Naidu, R.R.; Nawale, S.S.; Pawar, N.P.; Sharma, P.R.; Rajashree, R. A Study on Online Contract Signing Protocols. *Int. J. Eng. Adv. Technol.* **2014**, *4*, 188–190.
28. Louka, H.A. Necessity of parameter randomization in quantum contract signing. *Mat. Vesn.* **2017**, *69*, 65–74.

29. Feller, W. *An Introduction to Probability Theory and its Applications*; J. Wiley & Sons: New York, NY, USA, 1968; Volume 1.
30. Beeson, M.; Wiedijk, F. A Rigorous Theory of Infinite Limits. 2000. Available online: http://www.cs.kun.nl/~freek/notes/LimitTheory.pdf (accessed on 10 July 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Annals of Physics 401 (2019) 40–58

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: www.elsevier.com/locate/aop

Vanishing *k*-space fidelity and phase diagram's bulk–edge–bulk correspondence

P.D. Sacramento [a,*], B. Mera [a,b], N. Paunković [a,b,c]

[a] *CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*
[b] *Instituto de Telecomunicações, 1049-001 Lisboa, Portugal*
[c] *Departmento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

ARTICLE INFO

Article history:
Received 3 August 2018
Accepted 27 November 2018
Available online 4 December 2018

Keywords:
Fidelity
Topological transition

ABSTRACT

The fidelity between two infinitesimally close states and the fidelity susceptibility of a system are known to detect quantum phase transitions. Here we show that the *k*-space fidelity between two states far from each other and taken deep inside (bulk) of two phases, generically vanishes at the k-points where there are gapless points in the energy spectrum that give origin to the lines (edges) separating the phases in the phase diagram. We consider a general case of two-band models and present a sufficient condition for the existence of gapless points, given there are pairs of parameter points for which the fidelity between the corresponding states is zero. By presenting an explicit counter-example, we show that the sufficient condition is not necessary. Further, we show that, unless the set of parameter points is suitably constrained, the existence of gapless points generically implies accompanied pairs of parameter points with vanishing fidelity. Also, we show the connection between the vanishing fidelity and gapless points on a number of concrete examples (topological triplet superconductor, topological insulator, 1*d* Kitaev model of spinless fermions, BCS superconductor, Ising model in a transverse field, graphene and Haldane Chern insulator), as well as for the more general case of Dirac-like Hamiltonians. We also briefly discuss the relation between the vanishing fidelity and gapless points at finite temperatures.

© 2018 Elsevier Inc. All rights reserved.

* Corresponding author.
E-mail address: pdss@cefema.tecnico.ulisboa.pt (P.D. Sacramento).

https://doi.org/10.1016/j.aop.2018.11.021
0003-4916/© 2018 Elsevier Inc. All rights reserved.

1. Introduction

The fidelity and other quantum information signatures have been used to distinguish and characterize quantum phases, with particular emphasis on signaling their transitions [1,2]. Traditionally one compares states that differ infinitesimally due to some change of parameters of the Hamiltonian or due to some change in temperature or other intensive quantities associated with some reservoirs. The results together with some generalization, such as partial state fidelity [3,4] or the fidelity spectrum [5,6], have been used to detect quantum phase transitions including those of a topological nature. This includes topological insulators and topological superconductors [7–10].

The procedure was used to study the topological phases and transitions in various systems [11–38] and, in particular, in a two-dimensional triplet superconductor [39], which displays several trivial and topological phases, labeled by Chern numbers or a Z_2 invariant. Spinful electrons in the presence of a Zeeman term (that breaks time reversal symmetry) and in the presence of Rashba spin–orbit coupling are in a superconducting state with both singlet and triplet pairing symmetry (parity is broken due to the presence of the spin–orbit coupling). The Hamiltonian is written as

$$\begin{aligned}\hat{H} &= \frac{1}{2}\sum_{\boldsymbol{k}}\left(\boldsymbol{c}_{\boldsymbol{k}}^{\dagger},\boldsymbol{c}_{-\boldsymbol{k}}\right)\begin{pmatrix}\hat{H}_0(\boldsymbol{k}) & \hat{\Delta}(\boldsymbol{k})\\ \hat{\Delta}^{\dagger}(\boldsymbol{k}) & -\hat{H}_0^{T}(-\boldsymbol{k})\end{pmatrix}\begin{pmatrix}\boldsymbol{c}_{\boldsymbol{k}}\\ \boldsymbol{c}_{-\boldsymbol{k}}^{\dagger}\end{pmatrix}\\ &= \frac{1}{2}\sum_{\boldsymbol{k}}\left(\boldsymbol{c}_{\boldsymbol{k}}^{\dagger},\boldsymbol{c}_{-\boldsymbol{k}}\right)H_{\mathrm{BdG}}(\boldsymbol{k})\begin{pmatrix}\boldsymbol{c}_{\boldsymbol{k}}\\ \boldsymbol{c}_{-\boldsymbol{k}}^{\dagger}\end{pmatrix}\end{aligned} \tag{1}$$

where $\left(\boldsymbol{c}_{\boldsymbol{k}}^{\dagger},\boldsymbol{c}_{-\boldsymbol{k}}\right)=\left(c_{\boldsymbol{k}\uparrow}^{\dagger},c_{\boldsymbol{k}\downarrow}^{\dagger},c_{-\boldsymbol{k}\uparrow},c_{-\boldsymbol{k}\downarrow}\right)$,

$$\begin{aligned}\hat{H}_0 &= \epsilon_{\boldsymbol{k}}\sigma_0^s - M_z\sigma_z^s + \hat{H}_R\,,\\ \hat{H}_R &= \boldsymbol{s}\cdot\boldsymbol{\sigma}^{\boldsymbol{s}} = \alpha\left(\sin k_y\sigma_x^s - \sin k_x\sigma_y^s\right)\,,\end{aligned} \tag{2}$$

and $H_{\mathrm{BdG}}(\boldsymbol{k})$ is the so-called Bogoliubov–de Gennes Hamiltonian. Here, $\epsilon_{\boldsymbol{k}}=-2t(\cos k_x+\cos k_y)-\mu$ is the kinetic part, t denotes the hopping parameter set in the following as the energy scale, μ is the chemical potential, $\boldsymbol{k}$ is a wave vector in the xy plane, and we have taken the lattice constant to be unity. M_z is the Zeeman splitting term responsible for the magnetization, in energy units and the $\hat{H}_R$ is the Rashba spin–orbit term. α is measured in the energy units and $\boldsymbol{s}=\alpha(\sin k_y,-\sin k_x,0)$. The matrices $\sigma_x^s,\sigma_y^s,\sigma_z^s$ are the Pauli matrices acting on the spin sector, and σ_0^s is the 2×2 identity. The pairing matrix reads

$$\hat{\Delta}=i\left(\boldsymbol{d}\cdot\boldsymbol{\sigma}^s+\Delta_s\right)\sigma_y^s=\begin{pmatrix}-d_x+id_y & d_z+\Delta_s\\ d_z-\Delta_s & d_x+id_y\end{pmatrix}. \tag{3}$$

By introducing Pauli matrices acting in the Nambu space, τ_μ, $\mu=0,\ldots,3$, we can write the Bogoliubov–de Gennes Hamiltonian in the following compact form:

$$\begin{aligned}H_{\mathrm{BdG}}(\boldsymbol{k}) &= \epsilon_{\boldsymbol{k}}\tau_3\otimes\sigma_0^s - M_z\tau_3\otimes\sigma_3^s+\alpha\left(\sin(k_y)\tau_0\otimes\sigma_1^s-\sin(k_x)\tau_3\otimes\sigma_2^s\right)\\ &\quad+\Delta_t(\sin(k_x)\tau_2\otimes\sigma_0^s-\sin(k_y)\tau_1\otimes\sigma_3^s)-\Delta_s\tau_2\otimes\sigma_2^s.\end{aligned} \tag{4}$$

The system has a rich phase diagram with trivial and topological phases. These are shown in Fig. 1 considering $d_z=0$ and choosing $d_x=\Delta_t\sin k_y$, $d_y=-\Delta_t\sin k_x$. The Hamiltonian studied has therefore in general a 4×4 matrix structure. The problem is easily diagonalized and the lines where gapless points occur separate the different topological phases.

This model was studied in Refs. [40,41]. A particular interest was the study of entanglement and fidelity. These quantities were determined by numerical diagonalization of density matrices and the fidelity. Since the model factorizes in k-space the fidelity may be calculated for each momentum separately. In addition to the usual sensitivity of the fidelity around the critical points it was noted that some signature of these critical lines emerges at specific momentum values (associated with the points where the gap vanishes and the transitions occur) such that the k-space fidelity vanishes. This occurs even though it is calculated with density matrices that correspond to points in the phase

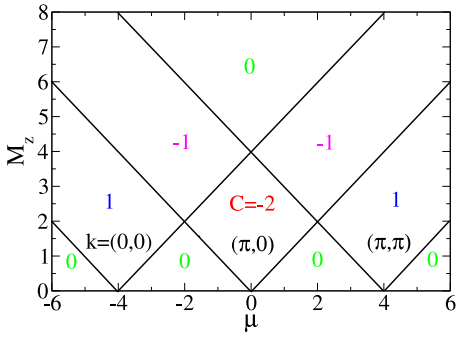


Fig. 1. (Color online) Phase diagram of a triplet superconductor as a function of chemical potential and Zeeman term. C is the Chern number. k is the momentum of each transition line.

diagram that are deep inside the various phases and not necessarily in the vicinity of the transition lines.

In this work we aim to understand better this result.

We begin by noting that the topology is not changed if $\Delta_s = 0, \alpha = 0$ as shown in [39]. If we take these values the 4×4 matrix decouples in two 2×2 matrices since the spin components do not get mixed anymore as

$$H_{\uparrow\uparrow} = \begin{pmatrix} \epsilon_{\boldsymbol{k}} - M_z & -i\Delta_t \left(\sin k_x - i \sin k_y\right) \\ i\Delta_t \left(\sin k_x + i \sin k_y\right) & -\epsilon_{\boldsymbol{k}} + M_z \end{pmatrix}, \tag{5}$$

and

$$H_{\downarrow\downarrow} = \begin{pmatrix} \epsilon_{\boldsymbol{k}} + M_z & -i\Delta_t \left(\sin k_x + i \sin k_y\right) \\ i\Delta_t \left(\sin k_x - i \sin k_y\right) & -\epsilon_{\boldsymbol{k}} - M_z \end{pmatrix}. \tag{6}$$

These matrices can be written in terms of Pauli matrices, $\boldsymbol{\sigma}$, as

$$\begin{aligned} H_{\uparrow\uparrow} &= (\epsilon_{\boldsymbol{k}} - M_z)\sigma_z - \Delta_t \sin k_y \sigma_x + \Delta_t \sin k_x \sigma_y , \\ H_{\downarrow\downarrow} &= (\epsilon_{\boldsymbol{k}} + M_z)\sigma_z + \Delta_t \sin k_y \sigma_x + \Delta_t \sin k_x \sigma_y . \end{aligned} \tag{7}$$

Denoting a vector $\boldsymbol{h}_{\sigma=\uparrow} = \boldsymbol{h}_{\uparrow\uparrow}$ and $\boldsymbol{h}_{\sigma=\downarrow} = \boldsymbol{h}_{\downarrow\downarrow}$ we get that the Hamiltonian matrices may be written in the form $\boldsymbol{h}_\sigma \cdot \boldsymbol{\sigma}$, with

$$\boldsymbol{h}_\uparrow = (-\Delta_t \sin k_y, \Delta_t \sin k_x, \epsilon_{\boldsymbol{k}} - M_z) \tag{8}$$

and

$$\boldsymbol{h}_\downarrow = (\Delta_t \sin k_y, \Delta_t \sin k_x, \epsilon_{\boldsymbol{k}} + M_z). \tag{9}$$

The reduction of the problem to two 2×2 matrices simplifies the problem considerably and an analytical solution for the fidelity is easy to obtain. Its analysis clarifies that the vanishing of the k-space fidelity at selected points is a general feature associated to a gapless point. We verify this result considering several models that display transitions either topological or non-topological.

In Section 2 we recall the definition of the fidelity and the k-space fidelity both in the finite temperature and zero temperature regimes. Then we apply it to the $2d$ triplet topological superconductor emphasizing the connection between the momenta where the fidelity vanishes and the spectrum gapless points at the transition lines. We perform an abstract analysis of the relation between zero-fidelity and gapless points, for the case of a general 2×2 Hamiltonian. Other models are considered also at zero temperature, both topological and non topological. In Section 3, models such as topological insulators, $1d$ Kitaev model of spinless fermions and the Haldane Chern insulator are discussed. Further, in Appendix A we analyze a conventional superconductor, the Ising model in a transverse field and graphene. Also, in Appendix B the triplet $2d$ superconductor is considered at

finite temperature showing that as temperature decreases the k-space fidelity approaches the regime of vanishing points at the gapless points that occur at the transition lines between different phases. In Section 4 we consider a generalization to higher dimensional Hamiltonians. The zero temperature fidelity is obtained and applied to a $3d$ topological insulator, further establishing the connection between transition lines with gapless points and zeros in the k-space fidelity. The reverse however is not always true. It is possible to find models where, although vanishing points in the fidelity can correspond to gapless excitations, they are not associated with transition lines. This is shown in Section 5 for a normal non-topological tight-binding model with a Zeeman term. The fidelity vanishes in extended regions that correspond to gapless points in the spectrum that are not associated with transition lines between phases. Another example is also considered that leads to a vanishing fidelity as a function of some control parameter introduced in a model of graphene that allows a continuous transition between the two opposite poles of the $\boldsymbol{h}$ vector Hamiltonian, with no specific gapless point in momentum space, since the spectrum vanishes for all momenta. We conclude with Section 6.

2. Fidelity

The quantum fidelity between two pure states (for two sets of parameters) is the absolute value of the overlap between the ground states for the two sets of parameters. In general, the quantum fidelity [42] between two states characterized by two density matrices ρ_1 and ρ_2, may be defined as the trace of the fidelity operator, $\mathcal{F}$, as $F(\rho_1,\rho_2) = \mathrm{Tr}\mathcal{F} = \mathrm{Tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}$. The fidelity operator $\mathcal{F}$ can be studied using different basis states, associated with different representations, such as position, momentum, energy or charge and spin.

Since the Hamiltonian is separable in momentum space, the density matrix operator for a momentum $\boldsymbol{k}$ may be defined as usual as

$$\hat{\rho}_{\boldsymbol{k}} = \frac{\mathrm{e}^{-\beta\widehat{H}_{\boldsymbol{k}}}}{Z_{\boldsymbol{k}}}, \tag{10}$$

In the diagonal basis it is written as

$$\rho_{\boldsymbol{k}} = \langle n|\hat{\rho}_{\boldsymbol{k}}|n\rangle = \frac{\mathrm{e}^{-\beta\langle n|\widehat{H}_{\boldsymbol{k}}|n\rangle}}{Z_{\boldsymbol{k}}}. \tag{11}$$

In Ref. [40] a basis representation for the density matrix in terms of the occupation numbers for a given momentum (and its symmetric) and the two spin projections was used. The eigenvalues of the density matrix are obtained if we diagonalize the Hamiltonian in the same basis. We considered the representation

$$\widetilde{H}_{\boldsymbol{k}} = \langle n_{\boldsymbol{k}_\uparrow} n_{-\boldsymbol{k}_\uparrow} n_{\boldsymbol{k}_\downarrow} n_{-\boldsymbol{k}_\downarrow} | \widehat{H}_{\boldsymbol{k}} | n_{\boldsymbol{k}_\uparrow} n_{-\boldsymbol{k}_\uparrow} n_{\boldsymbol{k}_\downarrow} n_{-\boldsymbol{k}_\downarrow}\rangle \tag{12}$$

The diagonalization of the Hamiltonian matrix in this enlarged basis is written as

$$\widetilde{H}_{\boldsymbol{k}}\mathbf{Q}_{\boldsymbol{k},n} = \lambda_{\boldsymbol{k},n}\mathbf{Q}_{\boldsymbol{k},n}\,, \quad n = 1,\ldots,16\,, \tag{13}$$

where $\mathbf{Q}_{\boldsymbol{k},n}$ is the eigenvector of $\widetilde{H}_{\boldsymbol{k}}$ associated to the eigenvalue $\lambda_{\boldsymbol{k},n}$. Note that n here is just an index number and should not be confused with the occupation number of Eq. (12). In the same basis the density matrix may be written as

$$\rho_{\boldsymbol{k}} = \frac{\mathrm{e}^{-\beta\widetilde{H}_{\boldsymbol{k}}}}{Z_{\boldsymbol{k}}}. \tag{14}$$

Therefore the eigenvalues of the density matrix may be written as $\rho_{\boldsymbol{k}}\mathbf{Q}_{\boldsymbol{k},n} = \Lambda_{\boldsymbol{k},n}\mathbf{Q}_{\boldsymbol{k},n}$ where

$$\Lambda_{\boldsymbol{k},n} = \frac{\mathrm{e}^{-\beta\lambda_{\boldsymbol{k},n}}}{\sum_{n'}\mathrm{e}^{-\beta\lambda_{\boldsymbol{k},n'}}}. \tag{15}$$

A simpler way to calculate the fidelity is to use a representation in the basis of the creation and destruction operators. In the case of the Sato and Fujimoto model this leads to a 4×4 representation (the same dimension of the Hamiltonian matrix).

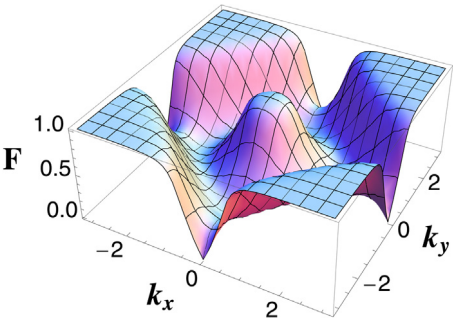


Fig. 2. (Color online) k-space fidelity for the 2d triplet superconductor with $\Delta_{t,1} = \Delta_{t,2} = 0.6$, $\mu_1 = -3$, $\mu_2 = -0.1$, $M_{z,1} = M_{z,2} = 0.5$, $T = 0$.

As mentioned in the introduction the problem may be further simplified noting that the density matrix for the Sato and Fujimoto model with $\Delta_s = \alpha = 0$ may be written as

$$\begin{aligned} \rho &= \rho_\uparrow \rho_\downarrow \\ \rho_\sigma &= \frac{\prod_{\boldsymbol{k}} e^{-\beta H^\sigma_{\boldsymbol{k}}}}{\prod_{\boldsymbol{k}} \mathrm{Tr}\left(e^{-\beta H^\sigma_{\boldsymbol{k}}}\right)} \end{aligned} \tag{16}$$

where now the matrices have a dimension 2×2.

The case of an Hamiltonian with a 2×2 structure has been considered before [43–45]. The k fidelity between two states ρ_1 and ρ_2 can be written as

$$F_{12}(\boldsymbol{k}) = \frac{2 + \sqrt{2\left(1 + A_{12}(\boldsymbol{k}) + B_{12}(\boldsymbol{k}) \boldsymbol{n}_{\boldsymbol{k},1} \cdot \boldsymbol{n}_{\boldsymbol{k},2}\right)}}{\sqrt{(2 + 2\cosh(\beta E_{\boldsymbol{k},1}/2))(2 + 2\cosh(\beta E_{\boldsymbol{k},2}/2))}} \tag{17}$$

where $\pm E_{\boldsymbol{k},i} = \pm|\boldsymbol{h}_{\boldsymbol{k},i}|$ are the energy eigenvalues of the Hamiltonian $H_{\boldsymbol{k},i} = \boldsymbol{h}_{\boldsymbol{k},i} \cdot \boldsymbol{\sigma}$, $\boldsymbol{n}_{\boldsymbol{k},i} = \boldsymbol{h}_{\boldsymbol{k},i}/|\boldsymbol{h}_{\boldsymbol{k},i}|$, $i = 1,2$, and

$$\begin{aligned} A_{12}(\boldsymbol{k}) &= \cosh(\beta E_{\boldsymbol{k},1}/2)\cosh(\beta E_{\boldsymbol{k},2}/2) \\ B_{12}(\boldsymbol{k}) &= \sinh(\beta E_{\boldsymbol{k},1}/2)\sinh(\beta E_{\boldsymbol{k},2}/2) \end{aligned} \tag{18}$$

In the zero temperature limit $\beta \to \infty$ the expression simplifies. The fidelity is given by

$$F_{12}(\boldsymbol{k}) = \sqrt{\frac{1}{2}\left(1 + \frac{\boldsymbol{h}_{\boldsymbol{k},1}}{|\boldsymbol{h}_{\boldsymbol{k},1}|} \cdot \frac{\boldsymbol{h}_{\boldsymbol{k},2}}{|\boldsymbol{h}_{\boldsymbol{k},2}|}\right)} \tag{19}$$

In this case $E_{\boldsymbol{k}} = |\boldsymbol{h}_{\boldsymbol{k}}|$. It is easy to see that if $\boldsymbol{h}_{\boldsymbol{k},1} = \boldsymbol{h}_{\boldsymbol{k},2}$ the fidelity is one.

2.1. *k-space fidelity of 2d triplet superconductor*

We begin by considering the case of zero temperature. Recalling that

$$F_{12} = F^\uparrow_{12} F^\downarrow_{12} \tag{20}$$

the fidelity for a given momentum may be obtained using Eqs. (8), (9).

In Fig. 2 we show the k-space fidelity for two density matrices that correspond to states in phases with different Chern numbers that are separated by a transition such that the spectrum gap closes at the point $\boldsymbol{k} = (\pi, 0)$ (and equivalent points). As obtained before numerically, the fidelity vanishes at these momenta values [40]. Similar results may be obtained for other examples, as shown in Ref. [40]. We also consider in the left panel of Fig. 3 two density matrices that correspond to a transition from

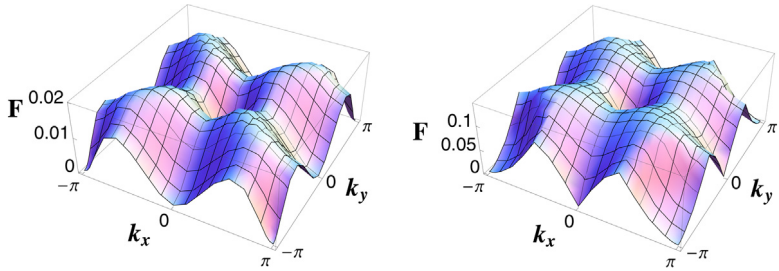


Fig. 3. (Color online) k-space fidelity for the $2d$ triplet superconductor with $\Delta_{t,1} = \Delta_{t,2} = 0.6$, (left panel) $\mu_1 = -6.0$, $\mu_2 = 6.0$, $M_{z,1} = M_{z,2} = 0.5$, $T = 0$. One gets the same result for $M_{z,1} = M_{z,2} = 0$. In the right panel $\mu_1 = -2.0$, $\mu_2 = 6.0$, $M_{z,1} = 4$, $M_{z,2} = 0.5$, $T = 0$. The behavior of the fidelity has linear dispersion around $\boldsymbol{k} = (0, 0), (\pi, 0)$ and quadratic around $\boldsymbol{k} = (\pi, \pi)$.

a trivial phase at $\mu_1 = -6$ to another trivial phase at $\mu_2 = 6$ and $M_{z,1} = M_{z,2} = 0.5$. In the right panel of the same figure we consider a density matrix at a topological phase with $C = -1$ to the same final trivial phase with $C = 0$. Tracing a straight line between the initial and final points in the phase diagram we see that in the left panel we cross twice gapless points at $\boldsymbol{k} = (0, 0), (\pi, 0), (\pi, \pi)$. In the case of the right panel we cross gapless points at $\boldsymbol{k} = (0, 0)$ (once), $\boldsymbol{k} = (\pi, \pi)$ (twice) and $\boldsymbol{k} = (\pi, 0)$ (once).

Taking the neighborhood of a point where the k-fidelity vanishes one can show that there is a factor proportional to the momentum displacement from the gapless point for each term that vanishes. Therefore evaluating the fidelity between two points in the phase diagram such as, for instance, $M_z = 0.5$ and $\mu_1 = -2$, $\mu_2 = 2$ there is factor proportional to $\boldsymbol{k}$ coming from each spin contribution leading to a factor of $\boldsymbol{k}^2$, and therefore a quadratic dispersion. In the case of Fig. 2 the dispersion is linear near $\boldsymbol{k} = (\pi, 0)$ since the gapless point is only crossed once and looking at Fig. 3 we the spectrum is quadratic near all gapless points in the left panel and linear near $\boldsymbol{k} = (0, 0), (\pi, 0)$ in the right panel and quadratic near $\boldsymbol{k} = (\pi, \pi)$. These results just confirm those obtained numerically before [40].

This result is explained next.

2.2. General result for 2×2 Hamiltonian matrix

Given a set $\mathcal{Q}$ of Hamiltonian parameters (which, in case of, say, effective Hamiltonians, may include temperature as well), for each momentum $\boldsymbol{k}$ we have Hamiltonians $H_q(\boldsymbol{k})$ and the corresponding Gibbs states $\rho_q(\boldsymbol{k}) = e^{-\beta H_q(\boldsymbol{k})}/Z(\boldsymbol{k})$, with $q \in \mathcal{Q}$ and β being the inverse temperature. Consequently, we consider the fidelity $F_{12}(\boldsymbol{k}) = F(\rho_{q_1}(\boldsymbol{k}), \rho_{q_2}(\boldsymbol{k}))$ and the energies $E_q(\boldsymbol{k})$.

At $\beta \to \infty$ limit, we are interested in finding the relation between the pairs of parameter points (q_1, q_2) for which the fidelity vanishes, $F_{12}(\boldsymbol{k}) = 0$, and the existence of critical gapless points q_c, for which $E_{q_c}(\boldsymbol{k}) = 0$.

Given the Hamiltonian

$$H_q(\boldsymbol{k}) = \boldsymbol{h}_q(\boldsymbol{k}) \cdot \boldsymbol{\sigma}, \tag{21}$$

its eigenvalues are $E_q^{\pm}(\boldsymbol{k}) = \pm|\boldsymbol{h}_q(\boldsymbol{k})|$, and the fidelity is

$$F_{12}(\boldsymbol{k}) = \sqrt{\frac{1}{2}\left(1 + \frac{\boldsymbol{h}_{q_1}(\boldsymbol{k})}{|\boldsymbol{h}_{q_1}(\boldsymbol{k})|} \cdot \frac{\boldsymbol{h}_{q_2}(\boldsymbol{k})}{|\boldsymbol{h}_{q_2}(\boldsymbol{k})|}\right)}. \tag{22}$$

Thus, gapless points are given by

$$E_{q_c}(\boldsymbol{k}) = |\boldsymbol{h}_{q_c}(\boldsymbol{k})| = 0, \tag{23}$$

and the condition that the fidelity vanishes translates to

$$\boldsymbol{h}_{q_1}(\boldsymbol{k}) \cdot \boldsymbol{h}_{q_2}(\boldsymbol{k}) = -|\boldsymbol{h}_{q_1}(\boldsymbol{k})||\boldsymbol{h}_{q_2}(\boldsymbol{k})|, \tag{24}$$

which implies that the angle between $\boldsymbol{h}_{q_1}(\boldsymbol{k})$ and $\boldsymbol{h}_{q_2}(\boldsymbol{k})$ is π.

This observation hints at the existence of the "gapless vector" $\boldsymbol{h}_{q_c}(\boldsymbol{k}) = \boldsymbol{0}$ between $\boldsymbol{h}_{q_1}(\boldsymbol{k})$ and $\boldsymbol{h}_{q_2}(\boldsymbol{k})$, defining the critical point q_c of potential quantum phase transition (see Section 5.1 for examples in which gapless excitations are not accompanied by transition lines). A simple sufficient condition that allows to infer gapless points (23) from the existence of zero fidelity pairs (24) is the linearity, with respect to q, of the function $\boldsymbol{h}_q(\boldsymbol{k})$, *provided* that the set of parameters $\mathcal{Q}$ is not too restricted. Indeed, condition (24) implies $\boldsymbol{h}_{q_2}(\boldsymbol{k}) = -\lambda \boldsymbol{h}_{q_1}(\boldsymbol{k})$, for some positive λ. Assuming that $\mathcal{Q}$ is a subspace of a real linear space, define $q_c = \mu q_1 + \nu q_2$, for some $\mu, \nu \in \mathbb{R}$. Assuming linearity, we have

$$\boldsymbol{h}_{q_c}(\boldsymbol{k}) = \boldsymbol{h}_{(\mu q_1 + \nu q_2)}(\boldsymbol{k}) = (\mu - \lambda\nu)\boldsymbol{h}_{q_1}(\boldsymbol{k}). \tag{25}$$

To satisfy (23), we need to satisfy

$$\mu = \lambda\nu, \tag{26}$$

which gives a line of critical points $q_c(\nu) = (\lambda q_1 + q_2)\nu$, parametrized by ν (note the above disclaimer — we require that for at least one $\nu \in \mathcal{Q}$, we also have $(\lambda q_1 + q_2)\nu \in \mathcal{Q}$). We show the above statement on the examples of topological insulator and 1*d* Kitaev model of spinless fermions (Sections 3.1 and 3.2, respectively), as well as for BCS superconductor, Ising model in a transverse field and graphene (Appendices A.1–A.3, respectively).

Nevertheless, linearity of $\boldsymbol{h}_q(\boldsymbol{k})$ with respect to q is not a necessary condition for the existence of gapless points. Consider $\boldsymbol{h}_{(\rho,\varphi)} = \rho\cos(\varphi)\boldsymbol{e_x} + \rho\sin(\varphi)\boldsymbol{e_y}$, with the parameters $\varphi \in [0, 2\pi)$, and $\rho \in [0, +\infty)$. Although $\boldsymbol{h}$ is not linear function of its parameter φ, it is obvious that $\boldsymbol{h}_{(\rho,\varphi)} \cdot \boldsymbol{h}_{(\rho,\varphi+\pi)} = -|\boldsymbol{h}_{(\rho,\varphi)}||\boldsymbol{h}_{(\rho,\varphi+\pi)}|$ and $E_{\rho=0,\varphi} = 0$ for every φ.

Note again that here, as in the above sufficient condition, we needed to assure the space of parameters $\mathcal{Q}$ to be "sufficiently large", in particular to include the $\rho = 0$ point. Had we restricted the parameters, say from $q = (\rho, \varphi)$ to $q = \varphi$, this simple "rotational" counter example shows that, at least in principle, there exist models for which there exists no $q = q_c$ for which $\boldsymbol{h}_{q_c}(\boldsymbol{k}) = \boldsymbol{0}$. Indeed, consider $\boldsymbol{h}_\varphi(\boldsymbol{k}) = \cos(\varphi)\boldsymbol{e_x} + \sin(\varphi)\boldsymbol{e_y}$, for $q = \varphi \in [0, 2\pi)$. We see that in this model there exist no gapless points, as the two energy bands are flat, $E_q^{\pm}(\boldsymbol{k}) = \pm 1$, while for each two "antipodal" parameter points φ and $\varphi + \pi$, we have that the corresponding fidelity is zero. See Section 5.2 for a particular tight binding model of graphene perturbed with a mass term.

One could pose an "opposite" question, whether the existence of a gapless point q_c, for which $E_{q_c}(\boldsymbol{k}) = |\boldsymbol{h}_{q_c}(\boldsymbol{k})| = 0$, implies the existence of pairs of parameters (q_1, q_2) for which the corresponding vectors $\boldsymbol{h}_{q_1}(\boldsymbol{k})$ and $\boldsymbol{h}_{q_2}(\boldsymbol{k})$ satisfy (24), for which the fidelity vanishes, $F_{12}(\boldsymbol{k}) = 0$. A simple counterexample shows that this, in general, is not the case. Take $q = (\rho, \varphi)$, such that $\rho \in [0, 1]$ and $\varphi \in [0, \pi/2]$. Define $\boldsymbol{h}_{(\rho,\varphi)} = \rho\cos(\varphi)\boldsymbol{e_x} + \rho\sin(\varphi)\boldsymbol{e_y}$. There is one gapless point, $\rho = 0$, but the fidelity is never zero, regardless of (q_1, q_2). Again, as in the above counter-examples, the way to avoid the existence of "zero-fidelity pairs" (q_1, q_2) is *to restrict*, this time the co-domain of the mapping $\boldsymbol{h}_q(\boldsymbol{k})$ to a set that excludes the existence of *any two* pairs of vectors for which $\boldsymbol{h}_{q_2}(\boldsymbol{k}) = -\lambda\boldsymbol{h}_{q_1}(\boldsymbol{k})$. Otherwise, having $\boldsymbol{h}_{q_c}(\boldsymbol{k}) = \boldsymbol{0}$, for some q_c, we can always find q_1 and q_2 for which

$$\begin{aligned} \boldsymbol{h}_{q_1}(\boldsymbol{k}) &= \boldsymbol{h}_{q_c}(\boldsymbol{k}) + \delta\boldsymbol{h}_{q_1}(\boldsymbol{k}) \\ \boldsymbol{h}_{q_2}(\boldsymbol{k}) &= \boldsymbol{h}_{q_c}(\boldsymbol{k}) + \delta\boldsymbol{h}_{q_2}(\boldsymbol{k}). \end{aligned} \tag{27}$$

Then, it follows that

$$\boldsymbol{h}_{q_1}(\boldsymbol{k}) \cdot \boldsymbol{h}_{q_2}(\boldsymbol{k}) = \delta\boldsymbol{h}_{q_1}(\boldsymbol{k}) \cdot \delta\boldsymbol{h}_{q_2}(\boldsymbol{k}). \tag{28}$$

Consider for instance the Sato and Fujimoto model simplified to the case where $\Delta_s = \alpha = 0$, since the topological properties are not changed. The transitions between the various phases occur at the momentum points $\boldsymbol{k} = (0, 0), (0, \pi), (\pi, \pi)$ (and equivalent points). Consider for instance the point $\boldsymbol{k} = (0, 0)$. The gapless point implies that

$$4t + \mu + M_z = 0 \tag{29}$$

The vanishing of the fidelity implies that

$$-1 = \mathrm{sgn}(4t_1 + \mu_1 + M_{z,1})\mathrm{sgn}(4t_2 + \mu_2 + M_{z,2}) \tag{30}$$

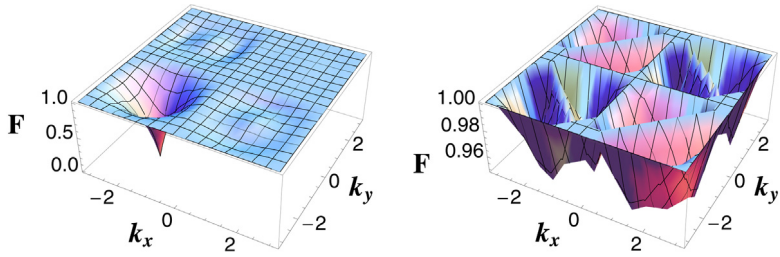


Fig. 4. (Color online) k-space fidelity for the topological insulator of Eq. (31) with (left panel) $t_{x,1,1} = t_{x,1,2} = t_{y,1,1} = t_{y,1,2} = 1$, $t_{2,1} = 1.2$, $t_{2,2} = 0.3$, $t'_{1,1} = t'_{1,2} = 0.5$, $\delta_1 = \delta_2 = 0.1$. In the right panel we take $t_{x,1,1} = t_{x,1,2} = 1$, $t_{y,1,1} = t_{y,1,2} = 0$, $t_{2,1} = 1.2$, $t_{2,2} = 0.3$, $t'_{1,1} = t'_{1,2} = 0$, $\delta_1 = \delta_2 = 0$.

which is satisfied if the signals are opposite. The transitions at the momentum origin occur in the vicinity of $\mu = -4$ if the magnetization is small (we fix $t = 1$). Similar expressions can be obtained in the vicinity of other transition lines.

3. Application to other systems

3.1. Topological insulator

A simple toy model for a two-dimensional topological insulator with two bands may be written as [46]

$$\begin{aligned} h_x &= \sqrt{2}t_{x,1}\left(\cos k_x + \cos k_y\right) \\ h_y &= \sqrt{2}t_{y,1}\left(\cos k_x - \cos k_y\right) \\ h_z &= 4t_2 \sin k_x \sin k_y + 2t'_1\left(\sin k_x + \sin k_y\right) + \delta \end{aligned} \tag{31}$$

The terms h_y and t'_1 break time reversal symmetry and the t_2 term breaks inversion symmetry. Since the system is two-dimensional, the system displays regimes with non-vanishing Chern numbers. For instance,

$$\begin{aligned} t_2 &> t'_1 - \frac{\delta}{4}\ ,\ C = 2 \\ t_2 &< t'_1 - \frac{\delta}{4}\ ,\ C = 1 \end{aligned} \tag{32}$$

At the points $\boldsymbol{k} = (\pm\pi/2, \pm\pi/2)$ both h_x and h_y vanish. Around these points and taking $\delta = 0$, h_z has the form $h_z = 4t_2 - 4t'_1$ at $(-\pi/2, -\pi/2)$, $h_z = 4t_2 + 4t'_1$ at $(\pi/2, \pi/2)$ and $h_z = -4t_2$ at the remaining points $(\pi/2, -\pi/2)$, $(-\pi/2, \pi/2)$. Therefore the momentum value that is associated with the transition from $t_2 > t'_1$ to $t_2 < t'_1$ is the one where the gap closes and the fidelity vanishes.

In Fig. 4 we show the k-space fidelity for this toy model. In the left panel we consider two phases such that phase 1 has $C = 2$ and phase 2 has $C = 1$, as discussed above. In the right panel we consider an example where there is no time reversal symmetry breaking. Also, there is no gapless point between the two sets of parameters. Therefore the fidelity has no zeros.

Another simple toy model that involves a transition between two topological regimes with $C = 2$ and $C = -2$ is the Hamiltonian

$$\begin{aligned} h_x &= \cos k_x + \cos k_y \\ h_y &= \cos k_x - \cos k_y \\ h_z &= t_2 \sin k_x \sin k_y \end{aligned} \tag{33}$$

The two regimes are obtained changing the sign of t_2. The gap closes at the points $\boldsymbol{k} = (\pi/2, \pi/2)$, $(\pi/2, -\pi/2)$, and equivalent points. This is clearly shown by the fidelity in Fig. 5.

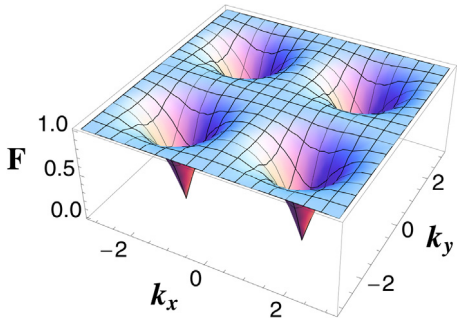


Fig. 5. (Color online) Fidelity for topological insulator of Eq. (33): $t_{2,1} = 1$, $t_{2,2} = -1$.

3.2. 1*d Kitaev model of spinless fermions*

In momentum space we may write the Kitaev model [47] as

$$\hat{H} = \frac{1}{2}\sum_k \left(c_k^\dagger, c_{-k}\right) \begin{pmatrix} \epsilon_k - \mu & -2i\Delta \sin k \\ 2i\Delta \sin k & -\epsilon_k + \mu \end{pmatrix} \begin{pmatrix} c_k \\ c_{-k}^\dagger \end{pmatrix} \tag{34}$$

with $\epsilon_k = -2t \cos k$. The Hamiltonian may be written using the Pauli matrices with

$$\boldsymbol{h} = (0, 2\Delta \sin k, \epsilon_k - \mu) \tag{35}$$

The eigenvalues are therefore $\pm|\boldsymbol{h}|$, where

$$|\boldsymbol{h}| = \sqrt{4\Delta^2 \sin^2 k + (-2t \cos k - \mu)^2} \tag{36}$$

The transitions lines occur for $\mu = 2$ and $k = \pi$, for $\mu = -2$ and $k = 0$ and for $\Delta = 0$ and $\cos k = -\mu/(2t)$. Therefore for $\Delta = 0$ and $\mu = 0$ the transition occurs at $k = \pi/2$. It is easy to check the vanishing of the fidelity. For this problem we can write that

$$\begin{aligned} \frac{\boldsymbol{h}_1 \cdot \boldsymbol{h}_2}{|\boldsymbol{h}_1||\boldsymbol{h}_2|} &= \left(4\Delta_1\Delta_2 \sin^2 k + (2t \cos k + \mu_1)(2t \cos k + \mu_2)\right) \\ &\times 1/\sqrt{4\Delta_1^2 \sin^2 k + (2t \cos k + \mu_1)^2} \\ &\times 1/\sqrt{4\Delta_2^2 \sin^2 k + (2t \cos k + \mu_2)^2} \end{aligned} \tag{37}$$

Considering for instance $\mu_1 = \mu_2 = 0$ it is easily seen that choosing for instance $\Delta_1 > 0$, $\Delta_2 < 0$ the expression reduces to -1 (vanishing fidelity) if $k = \pi/2$, which is the condition for the gapless point.

The vanishing of the fidelity may also be calculated directly using the eigenstates. At the points $\mu = 0$, $\Delta = \pm t$ the eigenvalues are ± 2 and the eigenvectors are

$$\psi_+ = \mathrm{sgn}\left[\cos\frac{k}{2}\right] \begin{pmatrix} -i\dfrac{\Delta}{t}\sin\dfrac{k}{2} \\ \cos\dfrac{k}{2} \end{pmatrix}, \tag{38}$$

$$\psi_- = \mathrm{sgn}\left[\cos\frac{k}{2}\right] \begin{pmatrix} \cos\dfrac{k}{2} \\ -i\dfrac{\Delta}{t}\sin\dfrac{k}{2} \end{pmatrix} \tag{39}$$

Taking now $\mu = 0$ but any value of Δ, the eigenvalues are

$$\lambda_\pm = \pm 2\sqrt{(t \cos k)^2 + (\Delta \sin k)^2} \tag{40}$$

The eigenvectors are (see for example Ref. [48])

$$\psi_+^{\Delta} = \begin{pmatrix} \dfrac{-i2\Delta \sin k}{\sqrt{2\lambda_+(\lambda_+ + 2t\cos k)}} \\ \sqrt{\dfrac{\lambda_+ + 2t\cos k}{2\lambda_+}} \end{pmatrix}, \tag{41}$$

$$\psi_-^{\Delta} = \begin{pmatrix} \sqrt{\dfrac{\lambda_- - 2t\cos k}{2\lambda_-}} \\ \dfrac{-i2\Delta \sin k}{\sqrt{2\lambda_-(\lambda_- - 2t\cos k)}} \end{pmatrix} \tag{42}$$

Consider for instance the states $\psi_+^{\Delta_1}$ and $\psi_+^{\Delta_2}$. Consider $\Delta_1 > 0, \Delta_2 < 0$. Their overlap is easily obtained

$$F_{12} = |-2\frac{|\Delta_1||\Delta_2| \sin^2 k}{\sqrt{\lambda_{+,1}(\lambda_{+,1} + 2t\cos k)\lambda_{+,2}(\lambda_{+,2} + 2t\cos k)}} + \sqrt{\frac{\lambda_{+,1} + 2t\cos k}{2\lambda_{+,1}}}\sqrt{\frac{\lambda_{+,2} + 2t\cos k}{2\lambda_{+,2}}}| \tag{43}$$

At the momentum $k = \pi/2$ we get that $\lambda_+ = 2|\Delta|$. Therefore the fidelity vanishes.

3.3. Haldane Chern insulator

We may also consider a graphene like model with the addition of hopping terms between nearest-neighbors on the same sublattice, t_2, with a periodic magnetic flux that breaks time-reversal inversal and therefore the possibility of a non-vanishing Chern number (but with zero total flux through a unit cell). This generalization was considered by Haldane [49] as an example of a topological Chern insulator in the absence of an external magnetic field. The magnetic flux is included by adding a phase to the hopping amplitude t_2. The Hamiltonian, including a mass term is given in momentum space by

$$\begin{aligned} H(\boldsymbol{k}) &= 2t_2 \cos\phi \sum_i \cos(\boldsymbol{k}\cdot\boldsymbol{b}_i) I \\ &+ t_1 \sum_i (\cos(\boldsymbol{k}\cdot\boldsymbol{a}_i)\sigma_1 + \sin(\boldsymbol{k}\cdot\boldsymbol{a}_i)\sigma_2) \\ &+ \left(M - 2t_2 \sin\phi \sum_i \sin(\boldsymbol{k}\cdot\boldsymbol{b}_i)\right)\sigma_3 \end{aligned} \tag{44}$$

Here t_1 is the hopping between nearest-neighbors between one sublattice and the other, M is the mass term and the lattice vectors are $\boldsymbol{a}_1 = (1, 0), \boldsymbol{a}_2 = (-1/2, \sqrt{3}/2), \boldsymbol{a}_3 = (-1/2, -\sqrt{3}/2)$ and $\boldsymbol{b}_1 = \boldsymbol{a}_2 - \boldsymbol{a}_3, \boldsymbol{b}_2 = \boldsymbol{a}_3 - \boldsymbol{a}_1, \boldsymbol{b}_3 = \boldsymbol{a}_1 - \boldsymbol{a}_2$. As shown in Fig. 6a, there are topological regions characterized by non-vanishing Chern numbers ± 1 for $|t_2/t_1| < 1/3$ that lead to non-vanishing Hall conductances. As a function of the phase ϕ and the gap magnitude, M, the non trivial phases occur if $|M/t_2| < 3\sqrt{3}|\sin\phi|$.

The fidelity may be calculated diagonalizing the Hamiltonian and by direct evaluation of the absolute value of the overlap of the eigenfunctions for two different sets of parameters. As for the above models, the k-space fidelity vanishes when comparing two states that are in two distinct phases that can be connected by a straight line in the phase diagram that cuts a transition or transition lines. As an example, we show in Fig. 6b the fidelity between states that differ by the value of $\phi = \pi/2, -\pi/2$ with the other parameters fixed at $t_1 = 1, t_2 = 0.25, M = 0.5$. The fidelity has zeros at the six corners of the Brillouin zone. This occurs because as one crosses from $\nu = 1$ to $\nu = 0$ and from $\nu = 0$ to $\nu = -1$, each transition line is characterized either by the Dirac zeros

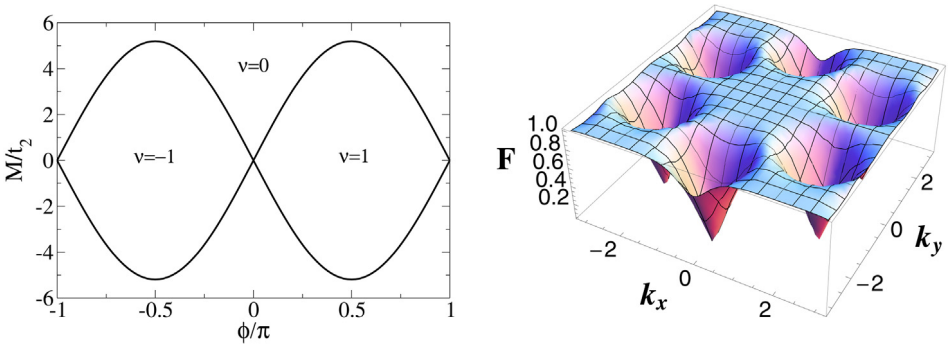


Fig. 6. (Color online) In the left panel we show the phase diagram of the Haldane model and in the right panel the fidelity with $t_1 = 1, t_2 = 0.25, M_1 = M_2 = 0.25, \phi_1 = \pi/2, \phi_2 = -\pi/2$.

at $\boldsymbol{K} = \frac{2\pi}{3}\left(1, \frac{1}{\sqrt{3}}\right)$ (and equivalent points) or $\boldsymbol{K}' = \frac{2\pi}{3}\left(1, -\frac{1}{\sqrt{3}}\right)$ (and equivalent points). Therefore the fidelity is linear around each of the vanishing points, as discussed above.

4. Generalization to higher dimensional hamiltonians

4.1. Fidelity

Consider a Hamiltonian of the form

$$H = \sum_{\mu=1}^{d} h^{\mu}\gamma_{\mu}, \tag{45}$$

where γ_μ, $\mu = 1, \ldots, d$ are Hermitian matrices corresponding to an irreducible representation of a Clifford algebra over the field of the complex numbers with d generators with Euclidean signature,

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu}I_{2^n}, \tag{46}$$

where I_{2^n} is the $2^n \times 2^n$ identity matrix and $n = \lfloor d/2 \rfloor$. From now on we will drop the index of the dimension of the vector space on the identity matrix. These matrices satisfy

$$\mathrm{Tr}\left(\gamma_\mu\gamma_\nu\right) = 2^n\delta_{\mu\nu}. \tag{47}$$

We have $H^2 = \|\boldsymbol{h}\|^2 I$ and, therefore, the eigenvalues are $\pm\|\boldsymbol{h}\|$, with $\|\boldsymbol{h}\|^2 \equiv \sum h^\mu h^\mu$. Let us assume $\boldsymbol{h} \neq 0$ and, without loss of generality, that $\|\boldsymbol{h}\| = 1$, i.e. $\boldsymbol{h} = (h^\mu)$ determines an element of the $(d-1)$-dimensional sphere S^{d-1}. Notice that

$$P = \frac{1}{2}\left(I - H\right), \tag{48}$$

commutes with H and is a projector. Similarly, $Q = I - P = (1/2)(I + H)$, also commutes with H and is a projector. Moreover,

$$H = Q - P = I - 2P. \tag{49}$$

So P corresponds to the projector onto the -1 eigenvalue sector and Q corresponds to the projector onto the $+1$ eigenvalue sector.

Associated to P we can build a density matrix

$$\rho(\boldsymbol{h}) = \frac{P}{\mathrm{Tr}\left(P\right)} = \frac{P}{2^{n-1}}. \tag{50}$$

The fidelity between two such density matrices, which we denote by $F(\boldsymbol{h}_1, \boldsymbol{h}_2)$, is given by

$$\begin{aligned} F(\boldsymbol{h}_1, \boldsymbol{h}_2) &= \mathrm{Tr}\Big(\sqrt{\sqrt{\rho(\boldsymbol{h}_1)}\rho(\boldsymbol{h}_2)\sqrt{\rho(\boldsymbol{h}_1)}}\Big) \\ &= \frac{1}{2^{n-1}}\mathrm{Tr}\Big(\sqrt{P_1P_2P_1}\Big), \end{aligned} \tag{51}$$

where we wrote $P_i = (1/2)(1 - \sum_\mu h_i^\mu \gamma_\mu) \equiv (1/2)(1 - H_i)$, $i = 1, 2$. Now, using $H_1H_2 + H_2H_1 = 2\langle \boldsymbol{h}_1, \boldsymbol{h}_2\rangle I$, with $\langle \boldsymbol{h}_1, \boldsymbol{h}_2\rangle = \sum_\mu h_1^\mu h_2^\mu$,

$$\begin{aligned} P_1P_2P_1 &= \frac{1}{8}\Big(2I - 2H_1 - H_2 + H_1H_2 + H_2H_1 - H_1H_2H_1\Big) \\ &= \frac{1}{8}\Big(2I - 2H_1 + 2\langle h_1, h_2\rangle(I - H_1)\Big) \\ &= \frac{1}{2}\big(1 + \langle \boldsymbol{h}_1, \boldsymbol{h}_2\rangle\big)P_1. \end{aligned}$$

So that,

$$F(\boldsymbol{h}_1, \boldsymbol{h}_2) = \sqrt{\frac{1}{2}\big(1 + \langle \boldsymbol{h}_1, \boldsymbol{h}_2\rangle\big)}, \tag{52}$$

similarly to the result in Eq. (22). Therefore if $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ form an angle of π, i.e., if they are antipodal with respect to each other, the fidelity will vanish.

4.2. Example: 3D topological insulator

Consider the following model for a 3D topological insulator [50,51]

$$H(\boldsymbol{k}) = v\tau^z\Big(\sum_\mu \sigma^\mu \sin(k_\mu)\Big) + \Big(M - t\sum_\mu \cos(k_\mu)\Big)\tau^x, \tag{53}$$

with $\mu = x, y, z$. The γ matrices,

$$\gamma_1 = \tau^z \otimes \sigma^x = \begin{pmatrix} \sigma^x & 0 \\ 0 & -\sigma^x \end{pmatrix},$$

$$\gamma_2 = \tau^z \otimes \sigma^y = \begin{pmatrix} \sigma^y & 0 \\ 0 & -\sigma^y \end{pmatrix},$$

$$\gamma_3 = \tau^z \otimes \sigma^z = \begin{pmatrix} \sigma^z & 0 \\ 0 & -\sigma^z \end{pmatrix},$$

$$\gamma_4 = \tau^x \otimes I_2 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix},$$

form an irreducible representation of a Clifford algebra in four generators with Euclidean signature. Our vector $\boldsymbol{h}$ is then given by

$$\boldsymbol{h}(\boldsymbol{k}) = \Big(v\sin(k_x), v\sin(k_y), v\sin(k_z), M - t\sum \cos(k_\mu)\Big). \tag{54}$$

The time-reversal operator is given by $\Theta = -i\big(I_2 \otimes \sigma^y\big)K$, where K is complex conjugation. Under time reversal $\boldsymbol{k} \to -\boldsymbol{k}$. The time-reversal invariant (TRI) momenta of the Brillouin zone B.Z. $\cong T^3$ are given by

$$(0,0,0),\ (\pi,0,0),\ (0,\pi,0),\ (0,0,\pi),$$
$$(\pi,\pi,0),\ (0,\pi,\pi),\ (\pi,0,\pi),\ (\pi,\pi,\pi).$$

The spatial inversion operator is given by $\Pi = \tau^x$. At a time-reversal invariant momentum $\boldsymbol{k}$, the Hamiltonian commutes with Π and also

$$H(\boldsymbol{k}) = \Big(M - t\sum_\mu \cos(k_\mu)\Big)\Pi \equiv m(\boldsymbol{k})\Pi. \tag{55}$$

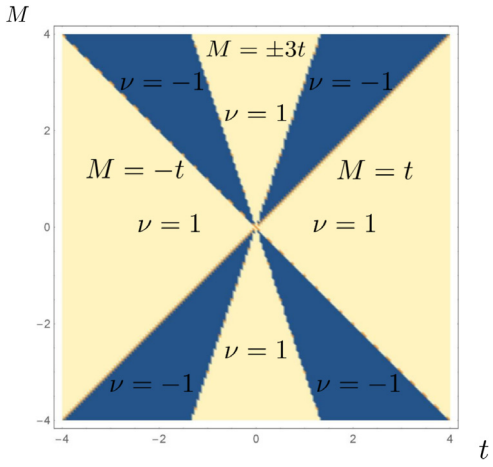


Fig. 7. The four lines $M = \pm t$ and $M = \pm 3t$ separate the phases where $\nu = \pm 1$.

The strong $\mathbb{Z}_2$ invariant is given by the product of the signs of the masses at TRI points,

$$\nu = \prod_{\{\boldsymbol{k}\in \text{B.Z.}:\ \boldsymbol{k}=-\boldsymbol{k}\}} \text{sgn}\big(m(\boldsymbol{k})\big) \in \mathbb{Z}_2. \tag{56}$$

Explicitly, it reads

$$\begin{aligned} \nu &= \text{sgn}\Big[(M-3t)(M-t)^3(M+t)^3(M+3t)\Big] \\ &= \text{sgn}\Big[(M^2-9t^2)(M^2-t^2)\Big]. \end{aligned} \tag{57}$$

The phase diagram is presented in Fig. 7.

We can now consider the k-space fidelity between groundstate subspaces associated with two Hamiltonians $H_1 \equiv H(M_1, t_2)$ and $H_2 \equiv H(M_2, t_2)$.

$$F(H_1(\boldsymbol{k}), H_2(\boldsymbol{k})) = \sqrt{\frac{1}{2}\Big(1 + \frac{\langle \boldsymbol{h}_1(\boldsymbol{k}), \boldsymbol{h}_2(\boldsymbol{k})\rangle}{\|\boldsymbol{h}_1(\boldsymbol{k})\|\,\|\boldsymbol{h}_2(\boldsymbol{k})\|}\Big)} \tag{58}$$

At TRI points we always have $H(\boldsymbol{k}) = m(\boldsymbol{k})\Pi$, or, equivalently, $\boldsymbol{h}(\boldsymbol{k}) = (0, 0, 0, m(\boldsymbol{k}))$. So for the fidelity to vanish at one point, we just need, for some TRI momenta $\boldsymbol{k}$

$$\text{sgn}(m_1(\boldsymbol{k})m_2(\boldsymbol{k})) = -1, \tag{59}$$

i.e., the masses have opposite signs. In fact, this condition means that the elements of the three-dimensional sphere S^3 defined by $\boldsymbol{h}_1(\boldsymbol{k})/|\boldsymbol{h}_1(\boldsymbol{k})|$ and $\boldsymbol{h}_2(\boldsymbol{k})/|\boldsymbol{h}_2(\boldsymbol{k})|$ are antipodal at this specific TRI momentum $\boldsymbol{k}$. More strikingly, at these TRI momenta, the fidelity will always be either zero or one. A straight line connecting $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ with this property will always close the gap at this TRI momentum.

For example, if we consider $v = 1$, take $M_1 = M_2 = 1$ and $t_1 = 1 \pm s$, with $s = 0.5$, we get at $\boldsymbol{k} = (0, 0, \pi)$,

$$\frac{\langle \boldsymbol{h}_1(\boldsymbol{k}), \boldsymbol{h}_2(\boldsymbol{k})\rangle}{\|\boldsymbol{h}_1(\boldsymbol{k})\|\,\|\boldsymbol{h}_2(\boldsymbol{k})\|} = -1. \tag{60}$$

In fact this holds for any of the following TRI points:

$$(0, 0, 0),\ (\pi, 0, 0),\ (0, \pi, 0),\ (0, 0, \pi),\ (\pi, \pi, \pi). \tag{61}$$

For $M = t$, the system is gapless at the TRI invariant momenta

$$(\pi, 0, 0),\ (0, \pi, 0),\ (0, 0, \pi). \tag{62}$$

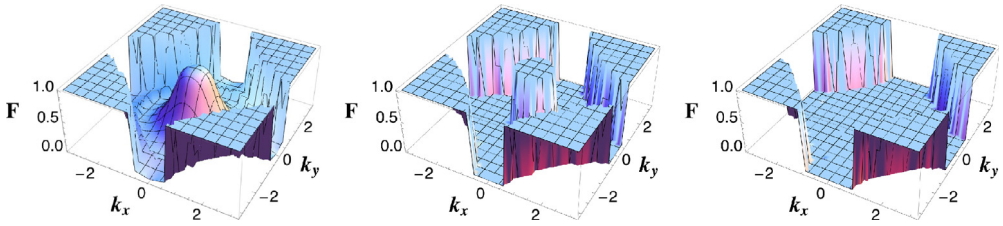


Fig. 8. (Color online) k-space fidelity for the $2d$ triplet superconductor with (left panel) $\Delta_{t,1} = 0.6$, $\Delta_{t,2} = 0.$, $\mu_1 = -3$, $\mu_2 = -0.1$, $M_{z,1} = M_{z,2} = 0.5$, $T = 0$, (middle panel) $\Delta_{t,1} = 0$, $\Delta_{t,2} = 0.$, $\mu_1 = -3$, $\mu_2 = -0.1$, $M_{z,1} = M_{z,2} = 0.5$, $T = 0$ and (right panel) $\Delta_{t,1} = 0$, $\Delta_{t,2} = 0.$, $\mu_1 = -3$, $\mu_2 = -0.1$, $M_{z,1} = M_{z,2} = 1.0$, $T = 0$.

The fact that the masses have opposite signs in an odd number of points implies there was a topological phase transition in between, as confirmed by the phase diagram.

5. Absence of transition lines and vanishing fidelity

As noted in Section 2.2, although we are primarily interested in the points of phase transitions, it is also interesting to analyze situations in which gapless points, in the presence of the vanishing fidelity, do not characterize some change of phase. Additionally, it is possible to find situations in which the fidelity vanishes, but this is not associated necessarily with a gapless point at some specific value of the coupling constants. We consider these two situations on the following two physical examples.

5.1. Gapless points without transitions

We now present an example of a physical model where there exist gapless points, see Eq. (23), which do not correspond to phase transition lines. Consider first the $2d$ triplet superconductor studied above. Turning off superconductivity, and since the normal term is not topological the topological nature is destroyed. In Fig. 8 we consider first (left panel) a transition from the superconductor to a point where superconductivity is turned off. The region where the fidelity vanishes widens but the overall features of Fig. 2 remain. However, turning off superconductivity altogether, as shown in the other panels of Fig. 8, the fidelity now vanishes in extended zones that correspond to gapless points in regions of momentum space and that are not associated with any transitions.

5.2. Zero fidelity without gapless points

A physical example where the fidelity goes to zero, see Eq. (24), at a given $\boldsymbol{k}$ and there is no parameter in the theory for which the gap closes at that same $\boldsymbol{k}$, see Eq. (23), is the following. Consider the tight binding of graphene. The Hamiltonian in momentum space is

$$H(\boldsymbol{k}) = tA(\boldsymbol{k})\sigma_+ + tA^*(\boldsymbol{k})\sigma_-, \tag{63}$$

where t is a hopping amplitude, σ are the pseudo-spin Pauli matrices and

$$A(\boldsymbol{k}) = \sum_{i=1}^{3} \exp(i\boldsymbol{k}\cdot\boldsymbol{a}_i), \tag{64}$$

where the $\boldsymbol{a}_i$ are nearest neighbor vectors. We can add a mass term, which amounts to taking

$$H(\boldsymbol{k}) \longrightarrow H(\boldsymbol{k}) + M\sigma_z. \tag{65}$$

Bring a parameter θ, such that $t = t(\theta) = t_0\cos(\theta)$ and $M = M(\theta) = t_0\sin(\theta)$. For $\theta = 0$, we have a vector $\boldsymbol{h}(\boldsymbol{k}) = (t_0\mathrm{Re}A(\boldsymbol{k}), -t_0\mathrm{Im}A(\boldsymbol{k}), 0)$ with gapless points at $\boldsymbol{K}$ and $\boldsymbol{K}'$. For $\theta = \pi$, the vector goes to $-h(\boldsymbol{k})$. Therefore, for every $\boldsymbol{k} \neq \boldsymbol{K}, \boldsymbol{K}'$, we always have,

$$F_{\boldsymbol{k}}(\theta_1 = 0, \theta_2 = \pi) = 0, \tag{66}$$

but

$$E(\boldsymbol{k}) = \pm t_0 \sqrt{\cos(\theta)^2 |A(\boldsymbol{k})|^2 + \sin(\theta)^2}, \tag{67}$$

will never be zero regardless of the value of θ (note that $t_0 \neq 0$ is fixed). Only for momenta $\boldsymbol{K}, \boldsymbol{K}'$ will the gap vanish, when $\theta = n\pi$, with n integer.

6. Conclusions

In this work we studied two-band models or more generally models that can be factorized to a set of two-bands. We investigated whether the k-space fidelity between states described by density matrices that correspond to points deep inside phases can provide information about the transition lines or sequence of transition lines that separate those phases. This extends previous numerical calculations for a $2d$ triplet spinful superconductor [40] where the result was identified.

In particular, we analyzed the relation between the existence of vanishing points of the k-space fidelity and gapless points. We analyzed general 2×2 Hamiltonians and presented a sufficient condition for the existence of gapless points, given there are pairs of parameter points for which the fidelity between the corresponding states is zero. By presenting an explicit counter-example, we showed that the sufficient condition is not necessary. Further, we showed that, unless the set of parameter points is suitably constrained, the existence of gapless points generically imply the accompanied pairs of parameter points with vanishing fidelity.

We showed explicitly that the vanishing fidelity is accompanied by the gapless points of zero-temperature quantum phase transitions on a number of concrete models: a topological insulator, the $1d$ Kitaev model of spinless fermions, the BCS superconductor, the Ising model in a transverse field, graphene and the Haldane model for a Chern insulator.

General Dirac-like Hamiltonians were also considered. We observed that the fidelity has the same form as in the two-band case. As a consequence, the same type of behavior is found, i.e., the k-space fidelity can vanish for points arbitrarily far from each other in parameter space, for momenta where the gap is found to close along a straight line joining the two points. As an example of this more general scenario, we considered a 3D topological insulator, classified by a $\mathbb{Z}_2$ topological invariant.

We also briefly discussed the finite-temperature case on the example of a $2d$ triplet superconductor.

Finally, we presented examples of systems in which, although vanishing fidelity *can* infer gapless points, those do not correspond to phase transition lines.

We conclude therefore that the results suggest that a vanishing fidelity strongly hints at a gapless point and eventually a transition between phases, but it does not hold in general and some specific counter-examples can be found. Then one can do the established procedure of going through the phase diagram step by step to search for a singular point.

Acknowledgments

The authors acknowledge discussions with Tharnier Puel de Oliveira, Pedro Ribeiro, Vítor Rocha Vieira, Nathan Goldman and Angelo Carollo. Partial support from FCT through grant UID/CTM/04540/2013 (Portugal) is acknowledged.

BM thanks the support from Fundação para a Ciência e Tecnologia (Portugal) namely through programmes PTDC/POPH/POCH and projects UID/EEA/50008/2013, IT/QuSim, IT/QuNet, ProQuNet, partially funded by EU FEDER (Portugal), from the EU FP7 project PAPETS (GA 323901) and from the JTF project NQuN, Portugal (ID 60478).

NP acknowledges the support of SQIG — Security and Quantum Information Group, the Instituto de Telecomunicações (IT) Research Unit (Portugal), ref. UID/EEA/50008/2013, the IT project QbigD funded by Fundação para a Ciência e Tecnologia (FCT) PEst-OE/EEI/LA0008/2013, and the FCT project Confident PTDC/EEI-CTP/4503/2014 (Portugal).

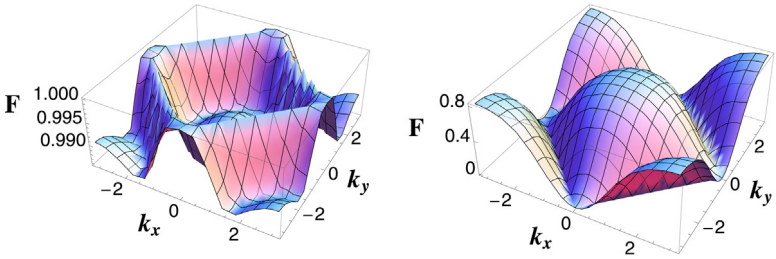


Fig. A.9. (Color online) BCS superconductor with $\Delta_1 = 1$, $\Delta_2 = 0.5$ and $\Delta_1 = 1$, $\Delta_2 = -1$.

Appendix A. Zero-temperature applications to other systems

A.1. BCS superconductor

Consider a conventional, non-topological, s-wave superconductor in two dimensions at finite temperature described by the effective mean-field BCS Hamiltonian

$$H_{BCS}^{eff} = \sum_{\boldsymbol{k}} \varepsilon_{\boldsymbol{k}}(n_{\boldsymbol{k}\uparrow} + n_{-\boldsymbol{k}\downarrow}) - \sum_{\boldsymbol{k}}(\Delta_{\boldsymbol{k}} c^{\dagger}_{\boldsymbol{k}\uparrow} c^{\dagger}_{-\boldsymbol{k}\downarrow} + \Delta^{*}_{\boldsymbol{k}} c_{-\boldsymbol{k}\downarrow} c_{\boldsymbol{k}\uparrow} - \Delta^{*}_{\boldsymbol{k}} \langle c_{-\boldsymbol{k}\downarrow} c_{\boldsymbol{k}\uparrow}\rangle), \tag{A.1}$$

To simplify we consider $\Delta_{\boldsymbol{k}} = \Delta$ a parameter independent of momentum but an extended s-wave superconductor could also be considered and it could also be determined self-consistently. We will be interested in situations where ρ_1 and ρ_2 correspond to points in parameter space, which we choose to be the temperature, T, and the gap, Δ, that are far apart and may be in the same or different thermodynamic phases [43].

In Fig. A.9 we consider a transition between two points at $\mu = 0$, one where the sign of Δ does not change and one where $\Delta_1 = -\Delta_2 = 1$. In the first case the fidelity is close to one, as expected since we are in the same phase. In the second case as Δ changes sign it crosses zero and there is a set of gapless points and the fidelity vanishes at those points since $\cos k_x + \cos k_y = 0$.

A.2. Ising model in a transverse field

The Ising model in a transverse field [52,53] described by

$$H = -\sum_{j=1}^{N} \left(\sigma_j^x \sigma_{j+1}^x + h\sigma_j^z\right) \tag{A.2}$$

where h is the transverse field, can be related to the Kitaev model performing a Jordan–Wigner transformation [54]. The fidelity between two states has been shown to be given by [1,2]

$$F(h, h') = \prod_{k\geq 0} \cos\left(\theta_k - \theta'_k\right) \tag{A.3}$$

where the Bogoliubov angles are defined (for two values of the transverse field) in the form

$$\cos(2\theta_k) = \frac{\cos k - h}{\sqrt{1 - 2h\cos k + h^2}} \qquad \sin(2\theta_k) = \frac{\sin k}{\sqrt{1 - 2h\cos k + h^2}} \tag{A.4}$$

The energy spectrum is given by

$$\epsilon_k = \sqrt{1 - 2h\cos k + h^2} \tag{A.5}$$

Taking $h = 1$ the spectrum becomes gapless at $k = 0$ and if $h = -1$ at $k = \pi$. Taking $k = 0$ we get that

$$\cos(2\theta_0) = \mathrm{sgn}(1-h) \tag{A.6}$$

Therefore

$$\begin{aligned}\cos\theta_0 &= \sqrt{\frac{1}{2}\left(1+\mathrm{sgn}(1-h)\right)}\\ \sin\theta_0 &= \sqrt{\frac{1}{2}\left(1-\mathrm{sgn}(1-h)\right)}\end{aligned} \tag{A.7}$$

Therefore

$$\begin{aligned}\cos\left(\theta_0-\theta_0'\right) = &\sqrt{\frac{1}{2}\left(1+\mathrm{sgn}(1-h)\right)}\sqrt{\frac{1}{2}\left(1+\mathrm{sgn}(1-h')\right)}\\ &+\sqrt{\frac{1}{2}\left(1-\mathrm{sgn}(1-h)\right)}\sqrt{\frac{1}{2}\left(1-\mathrm{sgn}(1-h')\right)}\end{aligned} \tag{A.8}$$

As a consequence if we choose two points on the h axis such that the $\mathrm{sgn}(1-h) = \mathrm{sgn}(1-h')$ the fidelity at $k = 0$ is one while if they are different the fidelity vanishes.

A.3. Graphene

We consider now a non-topological non-superconducting system such as graphene [55]. In order to gap the spectrum we add a mass term. We can model this massive graphene considering a mass term like $h_z = M$. In this case the system is non-topological since, even though there is a non-trivial Berry curvature emerging from each Dirac cone, the total Berry curvature cancels. However, introducing a mass term that depends on momentum such that it has opposite signs at the two Dirac cones leads to a non-vanishing Berry curvature and topological properties [49]. We consider therefore in addition the case

$$\begin{aligned}h_x &= 1+\cos\left(\sqrt{3}k_y\right)+\cos\left(\frac{\sqrt{3}}{2}k_y\right)\cos\left(\frac{3}{2}k_x\right)\\ &\quad-\sin\left(\frac{\sqrt{3}}{2}k_y\right)\sin\left(\frac{3}{2}k_x\right)\\ h_y &= \sin\left(\sqrt{3}k_y\right)+\sin\left(\frac{\sqrt{3}}{2}k_y\right)\cos\left(\frac{3}{2}k_x\right)\\ &\quad+\cos\left(\frac{\sqrt{3}}{2}k_y\right)\sin\left(\frac{3}{2}k_x\right)\\ h_z &= 4M\sin\left(\frac{\sqrt{3}}{2}k_y\right)\left(\cos\left(\frac{3}{2}k_x\right)-\cos\left(\frac{\sqrt{3}}{2}k_y\right)\right)\end{aligned} \tag{A.9}$$

The Dirac points are situated at $\boldsymbol{K} = \frac{2\pi}{3}\left(1, \frac{1}{\sqrt{3}}\right)$ and $\boldsymbol{K}' = \frac{2\pi}{3}\left(1, -\frac{1}{\sqrt{3}}\right)$ and $h_z(\boldsymbol{K}) = -h_z(\boldsymbol{K}')$.

In Fig. A.10 we consider the two cases where the mass is the same in both Dirac cones or it changes sign. We consider $M_1 = -M_2 = 0.5$. Both models show vanishing fidelity at the Dirac cones since by changing the sign of the mass at each Dirac point implies a crossing through zero energy.

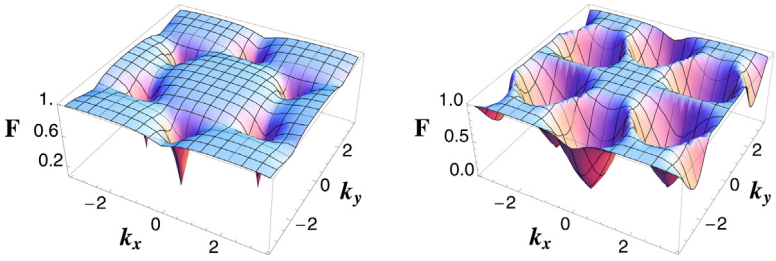


Fig. A.10. (Color online) Fidelity for graphene with $M_1 = -M_2 = 0.5$. In the left panel same mass on different Dirac cones and right panel opposite masses in different Dirac cones.

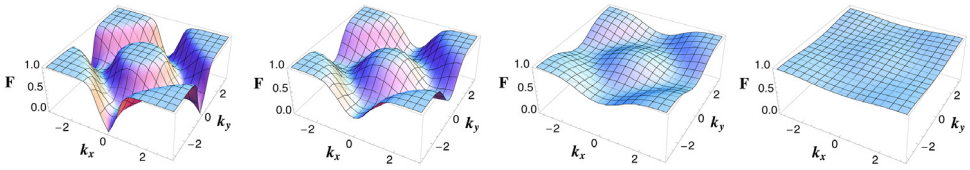


Fig. A.11. (Color online) Fidelity for the 2*d* triplet superconductor for $\Delta_{t,1} = \Delta_{t,2} = 0.6$, $\mu_1 = -2$, $\mu_2 = 0$, $M_{z,1} = M_{z,2} = 0.5$ and several temperatures: $T = 0, 0.25, 0.5, 1$.

Appendix B. Temperature effects on 2*d* triplet superconductor

The effect of a finite temperature leads to a smoothening of the fidelity and the vanishing points of the fidelity disappear. In Fig. A.11 we compare for the case of the Sato and Fujimoto model the k-space fidelity for different temperatures. Even though the vanishing points are absent, if the temperature is low there are signatures of their locations, as expected.

References

[1] P. Zanardi, N. Paunković, Phys. Rev. E 74 (2006) 031123.
[2] S.J. Gu, Internat. J. Modern Phys. B 24 (2010) 4371.
[3] N. Paunković, P.D. Sacramento, P. Nogueira, V.R. Vieira, V.K. Dugaev, Phys. Rev. A 77 (2008) 052302.
[4] H.Q. Zhou, arXiv:0704.2945.
[5] P.D. Sacramento, N. Paunković, V.R. Vieira, Phys. Rev. A 84 (2011) 062318.
[6] S.-J. Gu, W.-C. Wu, H.-Q. Lin, Internat. J. Modern Phys. B 27 (2013) 1350106.
[7] M.Z. Hasan, C.L. Kane, Rev. Modern Phys. 82 (2010) 3045.
[8] X.-L. Qi, S.-C. Zhang, Rev. Modern Phys. 83 (2011) 1057.
[9] J. Alicea, Rep. Progr. Phys. 75 (2012) 076501.
[10] A.P. Schnyder, S. Ryu, A. Furusaki, A.W.W. Ludwig, Phys. Rev. B 78 (2008) 195125.
[11] M.-F. Yang, Phys. Rev. B 76 (2007) 180403(R).
[12] S. Chen, L. Wang, Y. Hao, Y. Wang, Phys. Rev. A 77 (2008) 032111.
[13] S.-J. Gu, H.-M. Kwok, W.-Q. Ning, H.-Q. Lin, Phys. Rev. B 77 (2008) 245109.
[14] A. Hamma, W. Zhang, S. Haas, D.A. Lidar, Phys. Rev. B 77 (2008) 155111.
[15] D.F. Abasto, A. Hamma, P. Zanardi, Phys. Rev. A 78 (2008) 010301.
[16] S. Yang, S.-J. Gu, C.-P. Sun, H.-Q. Lin, Phys. Rev. A 78 (2008) 012304.
[17] L. Campus Venuti, M. Cozzini, P. Buonsante, F. Massel, N. Bray-Ali, P. Zanardi, Phys. Rev. B 78 (2008) 115410.
[18] D.F. Abasto, P. Zanardi, Phys. Rev. A 79 (2009) 012321.
[19] J.-H. Zhao, H.-Q. Zhou, Phys. Rev. B 80 (2009) 014403.
[20] E. Eriksson, H. Johannesson, Phys. Rev. A 79 (2009) 060301(R).
[21] M.M. Rams, B. Damski, Phys. Rev. Lett. 106 (2011) 055701.
[22] B. Damski, Phys. Rev. E 87 (2013) 052131.
[23] Q. Luo, J. Zhao, X. Wang, Phys. Rev. E 98 (2018) 022106.
[24] J. Sirker, M. Maiti, N.P. Konstantinidis, N. Sedlmayr, J. Stat. Mech: Theory Exp. (2014) P10032.
[25] E. König, A. Levchenko, N. Sedlmayr, Phys. Rev. B 93 (2016) 235160.
[26] C. Castelnovo, C. Chamon, Phys. Rev. B 77 (2008) 054433.

[27] S. Trebst, et al., Phys. Rev. Lett. 98 (2007) 070602.
[28] Z. Wang, T. Ma, S.-J. Gu, H.-Q. Lin, Phys. Rev. A 81 (2010) 062350.
[29] W.C. Yu, Y.C. Li, P.D. Sacramento, H.-Q. Lin, Phys. Rev. B 94 (2016) 245123.
[30] J. Eisert, M. Cramer, M.B. Plenio, Rev. Modern Phys. 82 (2010) 277.
[31] H.C. Jiang, Z. Wang, L. Balents, Nat. Phys. 8 (2012) 902.
[32] A. Kitaev, J. Preskill, Phys. Rev. Lett. 96 (2006) 110404.
[33] M. Levin, X.G. Wen, Phys. Rev. Lett. 96 (2006) 110405.
[34] H.C. Jiang, H. Yao, L. Balents, Phys. Rev. B 86 (2012) 024424.
[35] S. Furukawa, G. Misguich, Phys. Rev. B 75 (2007) 214407.
[36] S. Depenbrock, I.P. McCulloch, U. Schollwöck, Phys. Rev. Lett. 109 (2012) 067201.
[37] A.B. Kallin, M.B. Hastings, R.G. Melke, R.R.P. Singh, Phys. Rev. B 84 (2011) 165134.
[38] H. Yao, X.-L. Qi, Phys. Rev. Lett. 105 (2010) 080501.
[39] M. Sato, S. Fujimoto, Phys. Rev. B 79 (2009) 094504.
[40] T.P. Oliveira, P.D. Sacramento, Phys. Rev. B 89 (2014) 094512.
[41] T.P. Oliveira, P. Ribeiro, P.D. Sacramento, J. Phys.: Condens. Matter 26 (2014) 425702.
[42] R. Jozsa, J. Modern Opt. 41 (1994) 2315.
[43] N. Paunković, V.R. Vieira, Phys. Rev. E 77 (2008) 011129.
[44] B. Mera, C. Vlachou, N. Paunković, V.R. Vieira, Phys. Rev. Lett. 119 (2017) 015702, Supplemental Material.
[45] S. Tahir Amin, B. Mera, C. Vlachou, N. Paunković, V.R. Vieira, arXiv:1803.05021.
[46] M.A.N. Araújo, E.V. Castro, P.D. Sacramento, Phys. Rev. B 87 (2013) 085109.
[47] A.Y. Kitaev, Phys.-Usp. 44 (2001) 131.
[48] P.D. Sacramento, V.R. Vieira, Ann. Physics 391 (2018) 216.
[49] F.D.M. Haldane, Phys. Rev. Lett. 61 (1988) 2015.
[50] P. Hosur, S. Ryu, A. Vishwanath, Phys. Rev. B 81 (2010) 045120.
[51] R. Mong, V. Shivamoggi, Phys. Rev. B 83 (2011) 125109.
[52] E. Lieb, T. Schultz, D. Mattis, Ann. Physics 16 (1961) 407.
[53] B.K. Chakrabarti, A. Dutta, P. Sen, Quantum Ising Phases and Transitions in Transverse Ising Models, Springer-Verlag, 1996.
[54] P. Jordan, E. Wigner, Z. Phys. 47 (1928) 631;
Y.R. Wang, Phys. Rev. B 46 (1992) 151.
[55] A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, Rev. Modern Phys. 81 (2009) 109.

Fidelity and Uhlmann connection analysis of topological phase transitions in two dimensions

Syed Tahir Amin,[1,3,4,*] Bruno Mera,[1,3,4] Chrysoula Vlachou,[2,3,4] Nikola Paunković,[2,3,4] and Vítor R. Vieira[1,4]

[1]*Departamento de Física, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal*

[2]*Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal*

[3]*Instituto de Telecomunicações, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal*

[4]*CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal*

(Received 13 March 2018; revised manuscript received 29 October 2018; published 26 December 2018)

We study the behavior of the fidelity and the Uhlmann connection in two-dimensional systems of free fermions that exhibit nontrivial topological behavior. In particular, we use the fidelity and a quantity closely related to the Uhlmann factor in order to detect phase transitions at zero and finite temperature for topological insulators and superconductors. We show that at zero temperature both quantities predict quantum phase transitions: a sudden drop of fidelity indicates an abrupt change of the spectrum of the state, while the behavior of the Uhlmann connection signals equally rapid change in its eigenbasis. At finite temperature, the topological features are gradually smeared out, indicating the absence of finite-temperature phase transitions, which we further confirm by performing a detailed analysis of the edge states. Moreover, we performed both analytical and numerical analysis of the fidelity susceptibility in the thermodynamic limit, providing an explicit quantitative criterion for the existence of phase transitions. The critical behavior at zero temperature is further analyzed through the numerical computation of critical exponents.

DOI: 10.1103/PhysRevB.98.245141

I. INTRODUCTION

Topological phases of matter have received much attention in recent years due to their many potential applications in the emerging field of quantum technologies. Topological phases (TPs), in contrast to the Landau theory of phase transitions [1], are not described in terms of a local order parameter having a nontrivial value in the symmetry broken phase. Instead, they are described by global topological invariants [2–4], robust against continuous perturbations of the system preserving the symmetry and the spectral gap—no breaking of the symmetry occurs. These topological invariants, like Chern numbers [2], become observable through the response functions of the system coupled to an external field [2,5]. When terminating the system to the vacuum, there appear gapless modes which are exponentially localized at the boundary of the system, as predicted by the bulk-to-boundary principle [6]. On the other hand, it is impossible to deform a system from a topological phase with a given value of a topological invariant to the one where this invariant takes a different value without closing the gap. In this case, we say that a topological phase transition occurred, a particular example of a quantum phase transition [7]. From the quantum information theory point of view, there are two main approaches to the problem of quantum phase transitions. The first one is to study entanglement properties of the system [8–11]. The second approach is to study various distinguishability measures between quantum states, in particular the quantum fidelity [11–17], which we will consider here. When a system undergoes a phase transition, its ground state changes substantially and the quantum fidelity signals this change by a sudden drop in its value [14,15]. In this work, we extend the fidelity and Uhlmann connection analysis performed in [18,19] for the case of chiral topological phases in one-dimensional (1D) to 2D fermion topological phases, namely topological superconductors (TSCs) and insulators (TIs) classified by the first Chern number.

Our paper is organized as follows. In Sec. II we recall the notions of fidelity and Uhlmann connection and how their study can be used to probe phase transitions. In Sec. III we present the results of the fidelity and Uhlmann connection analysis for 2D TSCs and TIs. In Sec. IV, we analyze the noncommutativity of the thermodynamic and zero-temperature limits of the fidelity susceptibility and prove the absence of finite temperature phase transitions in the class of models considered. Moreover, we analyze the critical behavior at zero temperature by looking at critical exponents. In Sec. V we consider the system in a cylinder and study the edge states. Finally, in Sec. VI, we present our conclusions.

II. FIDELITY AND THE UHLMANN CONNECTION IN THE STUDY OF PHASE TRANSITIONS

In this section, we present the main concepts of the fidelity and the Uhlmann connection analysis of phase transitions. In particular, we will show the relationship between the fidelity and the quantity Δ associated to the Uhlmann factor, which we will in turn use to study phase transitions.

The general expression for the fidelity between two mixed states ρ and ρ' is

$$F(\rho,\rho') = \mathrm{Tr}\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}. \tag{1}$$

The set of mixed states is convex, but not linear in general, i.e., for any two mixed states ρ and ρ', and scalars λ and λ', the

*Corresponding author: tahiramin@tecnico.ulisboa.pt

 ©2018 American Physical Society

linear combination $\lambda\rho + \lambda'\rho'$ is not necessarily a mixed state. Nevertheless, a convex combination of ρ and ρ' is a mixed state, i.e., for $\lambda, \lambda' \geqslant 0$ and $\lambda + \lambda' = 1$, the linear combination $\lambda\rho + \lambda'\rho'$ belongs to the set of mixed states. This feature of nonlinearity imposes significant restrictions when it comes to performing a geometric study. On the other hand, we do not have this issue in the case of pure states, since a pure state $\rho = |\psi\rangle\langle\psi|$ can be treated as a projection on the subspace spanned by the state $|\psi\rangle$, therefore inheriting the geometric properties of the Hilbert space. Notice the $U(1)$-gauge freedom: $|\psi\rangle$ and $e^{i\phi}|\psi\rangle$ correspond to the same state $\rho = |\psi\rangle\langle\psi|$.

To overcome such restrictions, one can consider the concept of the purification of a mixed state, that is, any mixed state in a certain Hilbert space can be seen as the reduced state of a pure state in a different (larger) Hilbert space. Similarly, for each mixed state ρ in a Hilbert space H, we can consider the corresponding Hilbert-Schmidt space $H_w = H \otimes H^*$, where H^* is the dual of H, which contains the purifications w (in this case they are called amplitudes) such that $\rho = ww^\dagger$. Notice that there is a $U(n)$-gauge freedom in the choice of the amplitude, analogously to the $U(1)$-gauge freedom in the case of pure states: w and wU with U being unitary correspond to the same state $\rho = ww^\dagger$. In what follows, we will describe how a particular choice of the amplitude reveals the relationship between the fidelity and the Uhlmann connection. Two amplitudes w_1 and w_2, such that $\rho_1 = w_1 w_1^\dagger$ and $\rho_2 = w_2 w_2^\dagger$, are said to be parallel in the Uhlmann sense if they minimize the distance induced by the Hilbert-Schmidt inner product $\langle w_2, w_1\rangle = \mathrm{Tr}(w_2^\dagger w_1)$:

$$\begin{aligned} ||w_2 - w_1||^2 &= \mathrm{Tr}[(w_2 - w_1)^\dagger (w_2 - w_1)] \\ &= 2(1 - \mathrm{Re}\langle w_2, w_1\rangle). \end{aligned}$$

Minimizing $||w_2 - w_1||^2$ is equivalent to maximizing $\mathrm{Re}\langle w_2, w_1\rangle$:

$$\mathrm{Re}\langle w_2, w_1\rangle \leqslant |\langle w_2, w_1\rangle| = |\mathrm{Tr}(w_2^\dagger w_1)|.$$

Considering the polar decompositions $w_i = \sqrt{\rho_i}U_i, i \in \{1, 2\}$, where the U_i's are unitary matrices, the above inequality becomes

$$\mathrm{Re}\langle w_2, w_1\rangle \leqslant |\mathrm{Tr}(U_2^\dagger \sqrt{\rho_2}\sqrt{\rho_1} U_1)|.$$

Using the polar decomposition $\sqrt{\rho_2}\sqrt{\rho_1} = |\sqrt{\rho_2}\sqrt{\rho_1}|V$, with V unitary, and the cyclic property of the trace, we obtain

$$\mathrm{Re}\langle w_2, w_1\rangle \leqslant |\mathrm{Tr}(|\sqrt{\rho_2}\sqrt{\rho_1}|VU_1U_2^\dagger)|,$$

and the Cauchy-Schwarz inequality implies

$$\mathrm{Re}\langle w_2, w_1\rangle \leqslant \mathrm{Tr}|\sqrt{\rho_2}\sqrt{\rho_1}|,$$

with the equality holding for $VU_1U_2^\dagger = I$, where I is the identity [20]. Finally, we can write $|\sqrt{\rho_2}\sqrt{\rho_1}| = \sqrt{(\sqrt{\rho_2}\sqrt{\rho_1})^\dagger(\sqrt{\rho_2}\sqrt{\rho_1})}$ and get

$$\mathrm{Re}\langle w_2, w_1\rangle \leqslant \mathrm{Tr}\sqrt{(\sqrt{\rho_1}\rho_2\sqrt{\rho_1})} = F(\rho_1, \rho_2),$$

which shows the relationship between the fidelity and the Uhlmann connection, as characterized by V, the so-called Uhlmann factor. In other words, imposing two amplitudes, w_1 and w_2, to be parallel in the Uhlmann sense is equivalent to choosing, by means of the associated Uhlmann factor, the specific gauge $VU_1U_2^\dagger = I$, which gives rise to the fidelity between the states ρ_1 and ρ_2.

To illustrate this relationship better, let us consider the Uhlmann parallel transport condition. Suppose that $\rho(t)$ is a closed curve of density matrices parametrized by $t \in [0, 1]$. Then, given an initial state $\rho(0)$ and the corresponding amplitude $w(0)$, the Uhlmann parallel transport condition, taken for an infinitesimal time period δt, yields a unique curve in the space of amplitudes, along which $w(t)$ and $w(t + \delta t)$ are parallel, $\forall t$. In particular, the Uhlmann parallel transport operator is given as $V(t) = \mathcal{T}\exp\{-\int_0^t \mathcal{A}(d\rho/ds)ds\}$, where $\mathcal{T}$ is the time-ordering operator and $\mathcal{A}$ is the Uhlmann connection differential one-form. The Uhlmann factor is then $V = V(t + \delta t)V^\dagger(t)$. The length of the curve in the space of amplitudes—with respect to the metric induced by the Hilbert-Schmidt inner product—is equal to the length of the corresponding curve in the space of density matrices—with respect to the Bures metric. The respective Bures distance is given in terms of the fidelity as $d_B(\rho_1, \rho_2) = \sqrt{2[1 - F(\rho_1, \rho_2)]}$.

We proceed by presenting how we can apply the above concepts in the study of phase transitions. Consider two close points t and $t + \delta t$ in the parameter space and the respective states $\rho(t)$ and $\rho(t + \delta t)$ in the space of density matrices. If the two states belong to the same phase, they almost commute, hence $V \approx I$ and $\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)} \approx |\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)}|$. Moreover, since they are almost indistinguishable, $F(\rho(t), \rho(t + \delta t)) \approx 1$. On the other hand, if $\rho(t)$ and $\rho(t + \delta t)$ belong to different phases, they must be significantly different, thus their fidelity must be smaller than 1 [15]. The difference between the two states can be in their spectra or their eigenbases. In the case of the latter, we also have nontrivial $V \neq I$ [21].

To quantify the difference between the Uhlmann factor V and the identity, we will use the following quantity, as defined in [21]:

$$\begin{aligned} \Delta(\rho(t), \rho(t + \delta t)) := F(\rho(t), \rho(t + \delta t)) \\ - \mathrm{Tr}(\sqrt{\rho(t)}\sqrt{\rho(t + \delta t)}). \end{aligned} \quad (2)$$

Taking into account that $F(\rho(t), \rho(t + \delta t)) = \mathrm{Tr}(|\sqrt{\rho(t)}\sqrt{\rho(t + \delta t)}|)$ and the polar decomposition of $\sqrt{\rho(t)}\sqrt{\rho(t + \delta t)}$, as presented above, we can write

$$\Delta(\rho(t), \rho(t + \delta t)) = \mathrm{Tr}\{|\sqrt{\rho(t + \delta t)}\sqrt{\rho(t)}|(I - V)\}. \quad (3)$$

Indeed, $\Delta(\rho(t), \rho(t + \delta t))$ quantifies the difference between V and I; as for $V = I$ it is trivially zero.

The quantity Δ, just like the fidelity, is bounded by 0 and 1. To show this we recall the following theorem (see for example [22]):

$$|\mathrm{Tr}(AU)| \leqslant \mathrm{Tr}|A|, \quad (4)$$

for U being unitary and A any operator. Using the linearity of the trace and Eq. (3), we can write Δ as

$$\Delta(\rho, \rho') = \mathrm{Tr}\{|\sqrt{\rho}\sqrt{\rho'}|\} - \mathrm{Tr}\{|\sqrt{\rho}\sqrt{\rho'}|V\}.$$

From (4), by replacing $A = |\sqrt{\rho}\sqrt{\rho'}|$ and $U = V$, we obtain

$$|\mathrm{Tr}\{|\sqrt{\rho}\sqrt{\rho'}|V\}| \leqslant \mathrm{Tr}\{|\sqrt{\rho}\sqrt{\rho'}|\}.$$

Thus, we find the lower bound of $\Delta(\rho, \rho')$ to be

$$\Delta(\rho, \rho') \geqslant 0. \tag{5}$$

For the upper bound of $\Delta(\rho, \rho')$ we will use Eq. (2). Since $\mathrm{Tr}\{\sqrt{\rho}\sqrt{\rho'}\} \geqslant 0$, as $\sqrt{\rho}$ and $\sqrt{\rho'}$ are positive operators, $\Delta(\rho, \rho')$ is bounded from above by the maximum value of the fidelity:

$$\Delta(\rho, \rho') \leqslant 1. \tag{6}$$

Summarizing, from (5) and (6), we get

$$0 \leqslant \Delta(\rho, \rho') \leqslant 1. \tag{7}$$

For $\rho(t)$ and $\rho(t+\delta t)$ from the same phase, $F(\rho(t), \rho(t+\delta t)) = \mathrm{Tr}(\sqrt{\rho(t+\delta t)}\sqrt{\rho(t)}) \approx 1$, therefore $\Delta(\rho(t), \rho(t+\delta t)) \approx 0$, while for $\rho(t)$ and $\rho(t+\delta t)$ from different phases, $F(\rho(t), \rho(t+\delta t)) \neq 1$ and in the case where the Uhlmann factor is also nontrivial, we have $\Delta(\rho(t), \rho(t+\delta t)) \neq 0$.

In summary, the departure of fidelity from 1 and the departure of Δ from 0 are indicating the phase-transition points. In the particular case of topological insulators and superconductors, the respective phase transitions at zero temperature (topological quantum phase transitions) are featured by the closing of the energy gap, which occurs for certain critical values of the parameters of the Hamiltonian. The fidelity and Δ are expected to capture the gap-closing points, as the distinguishability of the states is enhanced close to the phase transition.

Regarding possible experimental observations of the fidelity and the Uhlmann connection, note that the fidelity induces the Bures metric on the space of density operators, which is nothing but the fidelity susceptibility. For large classes of orders, such as the symmetry-breaking orders of the Landau-Ginsburg type, it has been shown that the fidelity susceptibility is precisely the dynamical susceptibility of the system [21,23,24]. Thus, measuring the susceptibility directly reveals the fidelity-induced metric in such cases. Moreover, one can measure other directly observable quantities that are shown to be related to the fidelity and its metric; such is the case of tensor monopoles [25]. In connection to that, it has been shown that estimating various parameters is enhanced around the regions of criticality where the fidelity-induced metric (and the related Fisher information) exhibits singularities [26]. Thus, experiencing enhanced precision in parameter estimation is directly linked to the behavior of the fidelity; see for example a recent work [27].

As far as the Uhlmann connection is concerned, one can directly observe the associated phase, as in [28]. Alternatively, one can, as in the case of the fidelity, measure other directly observable physical quantities, such as dynamical susceptibilities and conductivities, that have been shown to directly depend on the mean Uhlmann curvature and the Uhlmann number introduced in [29].

III. FIDELITY AND Δ ANALYSIS OF 2D TOPOLOGICAL SYSTEMS

In this section we present our quantitative results for the fidelity and Δ analysis of 2D topological superconductors and insulators. In our study the parameters of the Hamiltonian and the temperature are treated on equal footing, that is, our parameter space consists of points (μ, T), where μ denotes the parameters of the Hamiltonian that drive the quantum phase transition and T denotes the temperature. We calculated the fidelity $F(\rho, \rho')$ and $\Delta(\rho, \rho')$ between the states $\rho(\mu, T)$ and $\rho' = \rho(\mu', T')$ with $\mu' = \mu + \delta\mu$ and $T' = T + \delta T$. We consider $\delta\mu$ and δT to be small, such that the points (μ, T) and (μ', T') of the parameter space, where we evaluate the fidelity and Δ, are close to each other. In particular, we treat the following two cases:

(i) $\delta\mu = 0.01$ and $\delta T = 0$, i.e., probing F and Δ only with respect to the parameters of the Hamiltonian, while keeping the temperature unchanged;

(ii) $\delta\mu = 0$ and $\delta T = 0.01$, i.e., probing F and Δ only with respect to the temperature, while keeping the parameters of the Hamiltonian unchanged.

The expressions for the aforementioned quantities are calculated with respect to the many-body thermal states,

$$\rho = e^{-\beta\mathcal{H}}/Z, \tag{8}$$

with the many-body Hamiltonian given by

$$\mathcal{H} = \tfrac{1}{2}\widetilde{\psi}^\dagger H_{\mathrm{BdG}}\widetilde{\psi}, \tag{9}$$

where $\widetilde{\psi}^\dagger = (\psi_1^\dagger \ldots \psi_N^\dagger \ \psi_1 \ldots \psi_N)$, in which, $\psi_i^\dagger$ and ψ_i are fermion creation and annihilation operators, N is the total number of degrees of freedom of the system, and

$$H_{\mathrm{BdG}} = \begin{pmatrix} H & \Delta \\ \Delta^\dagger & -H^T \end{pmatrix} \tag{10}$$

is the Bogoliubov–de Gennes Hamiltonian. Moreover, in the previous equation, H is an $N \times N$ Hermitian matrix and Δ is an $N \times N$ antisymmetric matrix (not to be confused with the Uhlmann quantity also denoted by Δ). Any free-fermion Hamiltonian having both charge-preserving terms of the form $\psi_i^\dagger\psi_j = \delta_{i,j} - \psi_j\psi_i^\dagger$ and superconducting terms of the form $\psi_i^\dagger\psi_j^\dagger$ or $\psi_i\psi_j$ can be cast in the form of $\mathcal{H}$, modulo a constant term. The Hamiltonian $\mathcal{H}$ can be diagonalized by a Bogoliubov-Valatin transformation. The occupation number, for each of the Bogoliubov quasiparticles, is of the Fermi-Dirac type

$$n(\varepsilon_i) = \frac{1}{e^{\beta\varepsilon_i} + 1},$$

where $i = 1, \ldots, N$ and ε_i are the energy eigenvalues. The expression for $\mathcal{H}$, in general, includes the chemical potential μ, and therefore it is included in the expression for ε_i. In the limit $T \to 0$, only the states which are below the Fermi level will be occupied, meaning $n(\varepsilon_i) \to 1$ for $\varepsilon_i < 0$ and $n(\varepsilon_i) \to 0$ for $\varepsilon_i > 0$. For the analytic derivation of the expressions of the fidelity and the Uhlmann quantity Δ, see Appendix A and the Supplemental Material of Ref. [5].

A. 2D topological superconductor

We consider the lattice Hamiltonian of a chiral p-wave superconductor [30–33],

$$\mathcal{H} = \sum_{ij}\Big(-t(c_{i+1,j}^\dagger c_{i,j} + c_{i,j+1}^\dagger c_{i,j}) - \frac{1}{2}(\mu - 4t)c_{i,j}^\dagger c_{i,j} + S(c_{i+1,j}^\dagger c_{i,j}^\dagger + i c_{i,j+1}^\dagger c_{i,j}^\dagger) + \mathrm{H.\,c.}\Big), \tag{11}$$

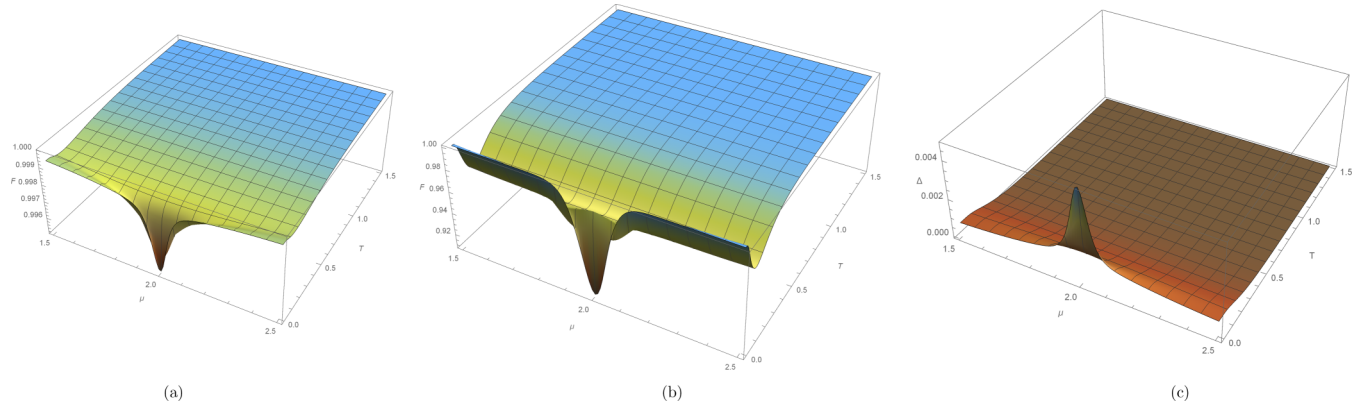


FIG. 1. Fidelity for the thermal state ρ, (a) when probing the parameter of Hamiltonian $\delta\mu = 0.01$, (b) when probing the temperature $\delta T = 0.01$, and (c) the quantity Δ when probing the Hamiltonian parameter $\delta\mu = 0.01$, for the 2D topological superconductor.

where $c_{i,j}$ and $c^\dagger_{i,j}$ are fermionic annihilation and creation operators for spinless fermions, respectively. S is the superconducting pairing, μ is the chemical potential, and t is the nearest-neighbor hopping. To simplify, we fix $|S| = t = \frac{1}{2}$. Since in real space the Hamiltonian is translationally invariant, we can perform a Fourier transformation to obtain

$$\mathcal{H} = \frac{1}{2}\sum_{\mathbf{k}} \widetilde{\psi}^\dagger_{\mathbf{k}} H_{\mathrm{BdG}}(\mathbf{k}) \widetilde{\psi}_{\mathbf{k}}, \tag{12}$$

where $\widetilde{\psi}^\dagger_{\mathbf{k}} = (c^\dagger_{\mathbf{k}}\ c_{-\mathbf{k}})$ and

$$H_{\mathrm{BdG}}(\mathbf{k}) = -\{\sin(k_y)\sigma_x + \sin(k_x)\sigma_y + [\mu - 2 + \cos(k_x) + \cos(k_y)]\sigma_z\}, \tag{13}$$

with $\sigma_{x,y,z}$ being the Pauli matrices. Finally, we apply a Bogoliubov-Valatin transformation to diagonalize $H_{\mathrm{BdG}}(\mathbf{k})$. At $T = 0$, the Chern number (Ch) characterizes two distinct nontrivial topological phases [34]: one with Ch $= 1$ for $0 < \mu < 2$ and one with Ch $= -1$ for $2 < \mu < 4$. For $\mu < 0$ and $\mu > 4$, the system is in a topologically trivial phase (Ch $= 0$).

In Fig. 1, we present the plots for the fidelity and Δ, when changing the parameters of the Hamiltonian and the temperature, respectively. We observe that at $T = 0$, in both cases the fidelity drops around the gap closing point at $\mu = 2$. As we increase the temperature these drops gradually disappear, indicating that the temperature smears out the topological features. Plot (c) shows the behavior of Δ, for $\delta\mu = 0.01$ and $\delta T = 0$, i.e., when we change the parameter of the Hamiltonian. At $T = 0$, Δ becomes nontrivial in the neighborhood of the gap-closing point at $\mu = 2$, signaling the topological phase transition, just like the fidelity. For higher T, this nontriviality is smeared out, indicating again the absence of finite-temperature phase transitions. By "smearing out" we mean that the fidelity and the Uhlmann quantity are smooth functions of temperature and they only reach their saturated values (1 and 0, respectively) in the $T \to \infty$ limit (in the same sense as it was previously used in [18,19,35]). For $\delta\mu = 0$ and $\delta T = 0.01$, i.e., when we only change the temperature, $\Delta = 0$ everywhere, as the Hamiltonians commute at different temperatures (we omitted the respective plot). The triviality of Δ, with respect to the change of temperature, can be explained by the fact that it is only sensitive to changes of the eigenbasis, while the fidelity is sensitive to both changes of the spectrum and eigenbasis [18].

B. 2D topological insulator

In two dimensions, a topological insulator can be realized by placing an atom with an internal degree of freedom at each site of a triangular lattice. This model is described by the following Hamiltonian [36]:

$$\mathcal{H} = \sum_{ij}[\psi^\dagger_{i+1,j}(t_1\sigma_x + it_3\sigma_z)\psi_{i,j} + \psi^\dagger_{i,j+1}(t_1\sigma_y + it_3\sigma_z)\psi_{i,j} + \psi^\dagger_{i+1,j+1}(t_2\sigma_z)\psi_{i,j} + \mathrm{H.\,c.}], \tag{14}$$

where $\psi^\dagger_{i,j}$ has components $c^\dagger_{\alpha,i,j}$, in which $\alpha = 1, 2$ labels the pseudospin.

Again, by taking advantage of the translational invariance of the Hamiltonian in real space, we transform it in momentum space, and considering periodic boundary conditions we obtain

$$H(\mathbf{k}) = 2t_1\cos(k_x)\sigma_x + 2t_1\cos(k_y)\sigma_y + \{2t_2\cos(k_x + k_y) + 2t_3[\sin(k_x) + \sin(k_y)]\}\sigma_z. \tag{15}$$

Notice here that $\mathcal{H}$ is charge symmetric and $H(\mathbf{k})$ is simply the Fourier transform of the block H in the matrix H_{BdG} from Eq. (10).

For simplicity, we fix $t_1 = t_3 = 1$, as in [34]. At $T = 0$, by varying the value of the parameter t_2, we can find four distinct topologically nontrivial phases with different Chern numbers: Ch $= \pm 2$ for $t_2 \lessgtr \mp 2$, and Ch $= \pm 1$ for $\mp 2 \lessgtr t_2 \lessgtr 0$. In Fig. 2, we show the respective plots for the fidelity and Δ. Similarly to the previous case of the 2D topological superconductor, at zero temperature we observe drops of the fidelity at the gap-closing points $t_2 = -2, 0, +2$, which signal the topological phase transitions [plots (a) and (b)]. When we increase the temperature, the fidelity goes gradually back to 1, indicating the absence of finite-temperature phase transitions.

Finally, in plot (c), we observe that Δ, when we vary the parameter of the Hamiltonian ($\delta t_2 = 0.01$, $\delta T = 0$), becomes nontrivial around the gap-closing points at $T = 0$, indicating the topological phase transitions, which are gradually smeared out as temperature increases. In the case where we only

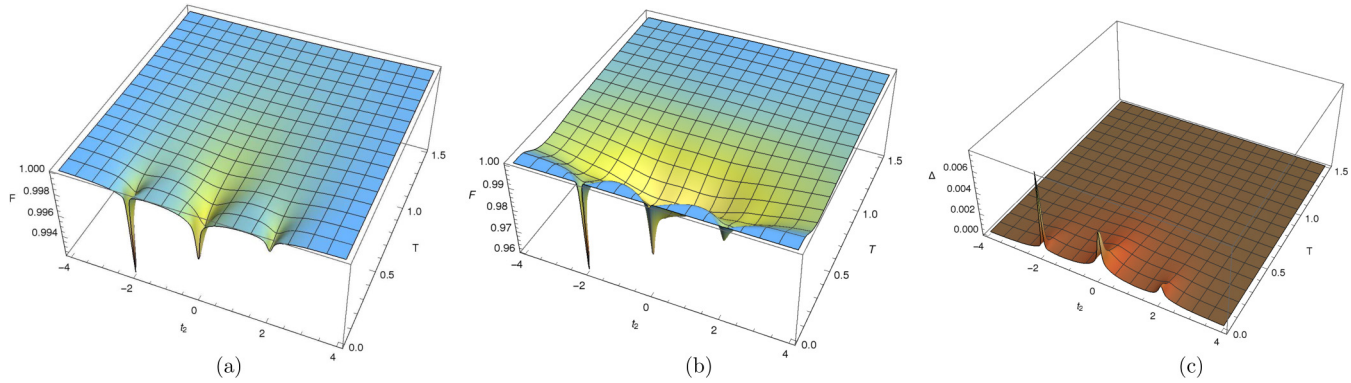


FIG. 2. Fidelity for the thermal state ρ, (a) when probing the parameter of Hamiltonian $\delta t_2 = 0.01$, (b) when probing the temperature $\delta T = 0.01$, and (c) the quantity Δ when probing the Hamiltonian parameter $\delta t_2 = 0.01$, for the 2D topological insulator.

change the temperature ($\delta t_2 = 0, \delta T = 0.01$) Δ is, again, zero everywhere, due to the fact the change of temperature only affects the spectrum of the Hamiltonian (we also omitted the respective plot).

Regarding the relation between the two quantities analyzed, note that the quantity Δ is manifestly trivial for the case of mutually commuting Hamiltonians. Thus, knowing the Hamiltonians do commute, $\Delta = 0$ brings no new information about the system's behavior. Since in our case Hamiltonians do not depend on the temperature, the behavior of Δ when only the temperature is varied tells us nothing about the existence of temperature-driven transitions, which are, at least in principle, possible even in those cases. Nevertheless, the fidelity, which is sensitive to *both* the change of Hamiltonian's eigenvectors, as well as its eigenvalues, can signal the existence of thermally driven phase transitions. Thus, one should consider the results of the fidelity as more general than those of the quantity Δ. Nevertheless, the possible nontrivial behavior of Δ gives additional and more refined information than the fidelity can provide.

IV. ABSENCE OF FINITE-TEMPERATURE PHASE TRANSITIONS: FIDELITY SUSCEPTIBILITY

In this section, we perform both analytical and numerical analysis of the fidelity susceptibility in the thermodynamic limit, providing an explicit quantitative criterion for the existence of phase transitions—if it diverges, there exists a phase transition, otherwise there is no phase transition. At the end, we also compute the respective critical exponents.

To fix notation, the Hamiltonian for the class of two-band systems considered can be written in terms of a vector $d^\mu(\mathbf{k}; M)$, where M is some parameter (e.g., a hopping amplitude) which drives the topological phase transition. More explicitly, for the topological superconductor, the many-body Hamiltonian is given by

$$\mathcal{H} = \frac{1}{2}\int_{\mathrm{BZ}} \frac{d^2k}{(2\pi)^2}\, \widetilde{\psi}^\dagger_{\mathbf{k}} d^\mu(\mathbf{k}; M)\sigma_\mu \widetilde{\psi}_{\mathbf{k}}, \tag{16}$$

where $\widetilde{\psi}^\dagger_{\mathbf{k}} = (c^\dagger_{\mathbf{k}}\ c_{-\mathbf{k}})$, σ_μ are the usual Pauli matrices and the Einstein summation convention is assumed. For the topological insulator, we have

$$\begin{aligned}\mathcal{H} &= \frac{1}{2}\int_{\mathrm{BZ}} \frac{d^2k}{(2\pi)^2}\, \widetilde{\psi}^\dagger_{\mathbf{k}} \begin{pmatrix} d^\mu(\mathbf{k}; M)\sigma_\mu & 0 \\ 0 & -[d^\mu(-\mathbf{k}; M)\sigma_\mu]^T \end{pmatrix} \widetilde{\psi}_{\mathbf{k}} \\ &= \int_{\mathrm{BZ}} \frac{d^2k}{(2\pi)^2}\, \psi^\dagger_{\mathbf{k}} d^\mu(\mathbf{k}; M)\sigma_\mu \psi_{\mathbf{k}}, \end{aligned} \tag{17}$$

where $\psi^\dagger_{\mathbf{k}} \equiv (c^\dagger_{1,\mathbf{k}}\ c^\dagger_{2,\mathbf{k}})$ is the Fourier transform of $\psi^\dagger_{i,j} = (c^\dagger_{1,i,j}\ c^\dagger_{2,i,j})$.

The absence of finite temperature phase transitions in the systems considered can be justified through a careful analysis of the fidelity susceptibility χ_{MM}. In doing so there are two limits to be considered: the thermodynamic limit, in which one considers the system on a finite lattice with periodic boundary conditions and then takes the number of its points to infinity, and the zero-temperature limit. The order in which one takes these two limits is important and, in fact, it produces different results. The fidelity susceptibility is defined through the formula

$$\frac{1}{2}\chi_{MM}\delta M^2 = -\frac{\ln[F(\rho(M), \rho(M+\delta M))]}{N^2}, \tag{18}$$

where the fidelity is taken with respect to infinitesimally close thermal states, of a fixed temperature, $\rho(M)$ and $\rho(M+\delta M)$, with respect to a variation of the parameter M, namely $M \to M + \delta M$, and N^2 is the number of lattice sites (N sites on each of the two directions in real space). It is the pullback by $\rho : M \mapsto \rho(M)$ of the Bures metric (appropriately normalized to take the thermodynamic limit). Taking the zero-temperature limit and then the thermodynamic limit yields the following expression for the fidelity susceptibility of the topological insulator:

$$\chi_{MM} = \int_{\mathrm{BZ}} \frac{d^2k}{(2\pi)^2}\, \frac{1}{4}\delta_{\mu\nu}\, \frac{\partial n^\mu}{\partial M}\, \frac{\partial n^\nu}{\partial M}, \tag{19}$$

where $n^\mu = d^\mu/|d|$ is the normalized Bloch vector. The previous formula manifests the pure state character of the zero-temperature limit of the thermal state. Indeed, it is the integral over the first Brillouin zone of the pullback of the Fubini-Study metric on the Bloch sphere by family of maps $M \mapsto n^\mu(M) = d^\mu(M)/|d(M)|$ parametrized by the quasimomentum $\mathbf{k} \in \mathrm{BZ}$. The above integral

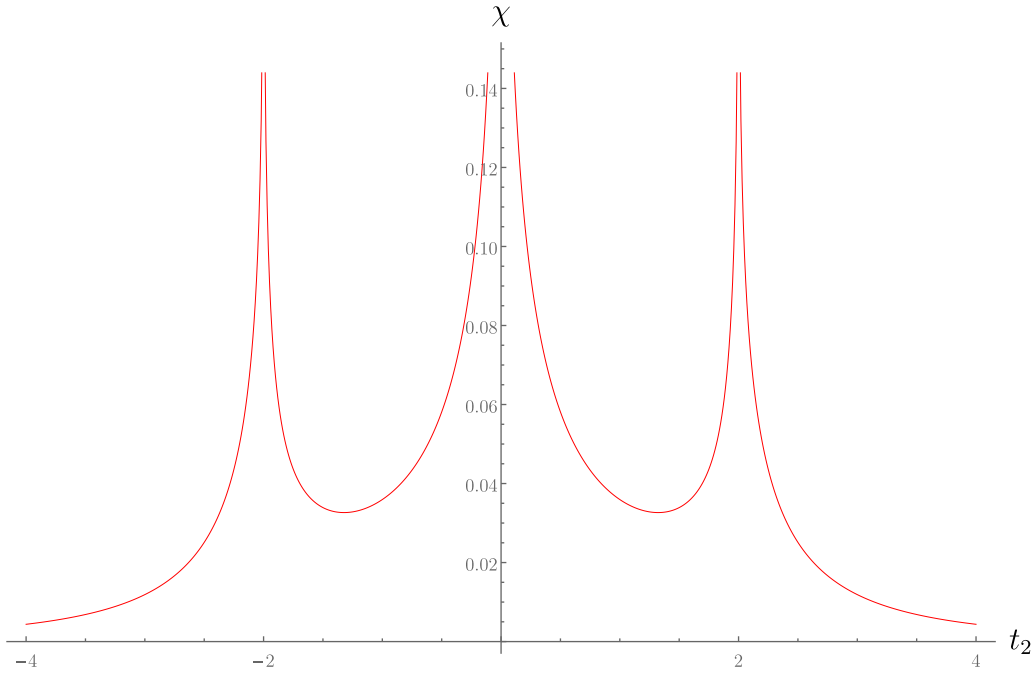


FIG. 3. The thermodynamic limit of the fidelity susceptibility at zero temperature for the 2D topological insulator model. The singularities appear precisely at the critical points of quantum phase transition where the gap closes.

can be numerically evaluated, and the results for the case of topological insulator, given by the Hamiltonian (14), are presented in Fig. 3. The fidelity susceptibility diverges precisely at the gapless points that define quantum criticality, and are given by $M \equiv t_2 = 0, \pm 2$. By taking finite temperature expression for the fidelity susceptibility, and performing the thermodynamic limit, we obtain

$$\chi_{MM} = \int_{\mathrm{BZ}} \frac{d^2k}{(2\pi)^2} \left\{ \frac{1}{4} \left[\frac{\cosh[\beta E(\mathbf{k})] - 1}{\cosh[\beta E(\mathbf{k})]} \delta_{\mu\nu} \frac{\partial n^\mu}{\partial M} \frac{\partial n^\nu}{\partial M} + \frac{\beta^2}{\cosh[\beta E(\mathbf{k})] + 1} \left(\frac{\partial E(\mathbf{k})}{\partial M} \right)^2 \right] \right\}, \tag{20}$$

where $E(\mathbf{k}) = |d(\mathbf{k})|$ and $\beta = 1/T$ is the inverse of the temperature (see Appendix C for the derivation). One is then interested in taking the limit $T \to 0$ of the previous expression. A naive approach would give the same as Eq. (19). However, if one takes into consideration the existence of the gapless points, the result is different. While the expression of Eq. (19) yields singularities at the critical points, the later Eq. (20) is perfectly regular. The reason is the second term coming from the variation of the energy which cures the divergence. When starting with the Boltzmann-Gibbs states, taking the thermodynamic limit and then taking the zero-temperature limit of the fidelity susceptibility, the second term appears and is relevant, especially in the neighbourhood of gap closing points. This term is precisely the difference between taking the zero-temperature limit and then the thermodynamic limit, and vice versa. For generic nonzero values of $E(\mathbf{k})$, the limit $\beta \to +\infty$ of the integrand function yields

$$\frac{1}{4} \delta_{\mu\nu} \frac{\partial n^\mu}{\partial M} \frac{\partial n^\nu}{\partial M}, \tag{21}$$

as expected. However, near the gap closing points this is not true. In this case, we take β finite and start by expanding in $E(\mathbf{k})$,

$$\begin{aligned} &\frac{1}{8} [\beta E(\mathbf{k})]^2 \delta_{\mu\nu} \frac{\partial n^\mu}{\partial M} \frac{\partial n^\nu}{\partial M} + \frac{1}{8} \beta^2 \left(\frac{\partial E(\mathbf{k})}{\partial M} \right)^2 + O\{[\beta E(\mathbf{k})]^4\} \\ &= \frac{1}{8} \beta^2 \delta_{\mu\nu} \frac{\partial (E n^\mu)}{\partial M} \frac{\partial (E n^\nu)}{\partial M} + O\{[\beta E(\mathbf{k})]^4\} \\ &= \frac{1}{8} \beta^2 \delta_{\mu\nu} \frac{\partial d^\mu}{\partial M} \frac{\partial d^\nu}{\partial M} + O\{[\beta E(\mathbf{k})]^4\}. \end{aligned}$$

The last expression is too coarse to take the zero-temperature limit. Notice that we lost the denominators in this approximation. The denominators in both terms are roughly the same, both in the $\beta \to +\infty$, while keeping $E(\mathbf{k})$ finite, but nonzero and $E(\mathbf{k}) \to 0$, while keeping β finite. We have the inequality

$$\frac{1}{2\cosh(x)} < \frac{1}{\cosh(x) + 1} < \frac{1}{\cosh(x)}, \quad \text{for all } x \in \mathbb{R}. \tag{22}$$

Therefore, we can safely upper bound the integrand by

$$\frac{1}{4} \frac{\beta^2}{\cosh[\beta E(\mathbf{k})]} \delta_{\mu\nu} \frac{\partial d^\mu}{\partial M} \frac{\partial d^\nu}{\partial M} + O\{[\beta E(\mathbf{k})]^4\} \tag{23}$$

in a neighborhood of the gap closing points. Notice that the integrand will be positive, or zero, since it is a pullback of a Riemannian metric. However, if we now take the zero-temperature limit, this upper bound vanishes. Since the singularities at zero temperature were coming from the gap closing points, this shows why there are no finite temperature phase transitions in the models discussed.

The above analysis of the fidelity susceptibility was performed for the case of the topological insulator. The case of the topological superconductor is identical, apart from a constant factor of $1/2$; see Appendix A for details.

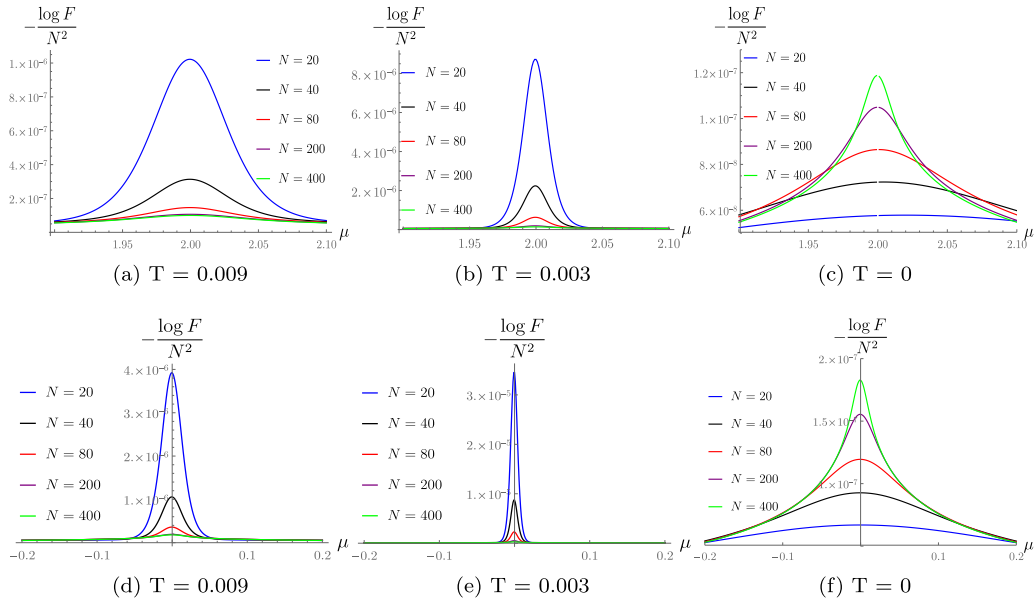


FIG. 4. $-\ln F/N^2$ for different system size against parameter that control quantum phase transition. Panels (a)–(c) are for 2D topological superconductor, while panels (d)–(f) are for 2D topological insulator.

The analytical study is further confirmed by the numerical evaluation of $-\ln F/N^2$ for both the topological superconductor and insulator, presented in Fig. 4.

The absence of finite temperature phase transitions means that at temperatures different from zero, the fidelity susceptibility χ does not signal phase transitions, as it does for temperature zero. By changing the parameters of the Hamiltonian at $T = 0$, we are able to close the energy gap and have a phase transition from a topological to a trivial phase and/or vice versa, which is detected by the fidelity susceptibility, due to the enhanced distinguishability of the states close to the point of the phase transition. For finite temperatures ($T > 0$), however, we do not observe any changes of χ, which implies the nonexistence of phase transitions, driven neither by the parameters of the Hamiltonian nor by the temperature. In this context, one can talk about *a crossover from the zero-temperature topological phase to the infinite-temperature nontopological phase*, since on one hand, already at $T = 0$, we have topological and trivial phases, and on the other hand, the fidelity susceptibility detects the phase transitions between them, which do not exist for $T > 0$.

Finally, one can perform the asymptotic analysis of the fidelity susceptibility in the vicinity of the critical points to obtain the critical exponents. For the topological superconductor, we performed a least-squares fit to the susceptibility to a power law $\chi \propto M^{-b}$ in a neighborhood of $M \equiv \mu = 0$ and found $b = 0.22 \pm 0.0036$. Furthermore, we also performed a least-squares fit to the susceptibility to a power law $\chi \propto (M-2)^{-b}$ in a neighborhood of $M \equiv \mu = 2$ and found $b = 0.26 \pm 0.0079$; see Fig. 5 for both cases. Analogously, for the topological insulator, we performed a least-squares fit to the susceptibility to a power law $\chi \propto M^{-b}$ in a neighborhood of $M \equiv t_2 = 0$ and found $b = 0.33 \pm 0.012$. Additionally, we also performed a least-squares fit to the susceptibility to a power law $\chi \propto (M-2)^{-b}$ in a neighborhood of $M \equiv t_2 = 2$ and found $b = 0.33 \pm 0.012$; see Fig. 6 for both cases.

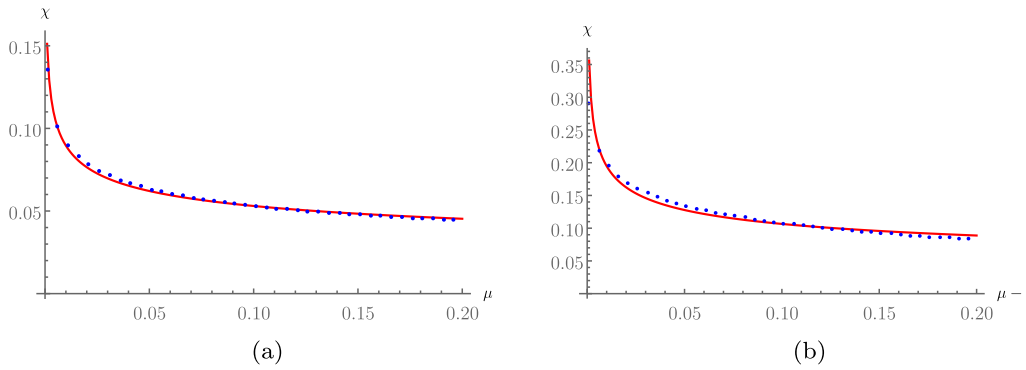


FIG. 5. The fidelity susceptibility χ_{MM} for the topological superconductor, as a function of $(M - M_c)$. The points are the result of numerical integration. The red curve is the result of the fit. (a) The fidelity susceptibility χ_{MM} as a function of M in a neighborhood of $M \equiv \mu = 0$. (b) The fidelity susceptibility χ_{MM} as a function of $(M-2)$ in a neighborhood of $M \equiv \mu = 2$.

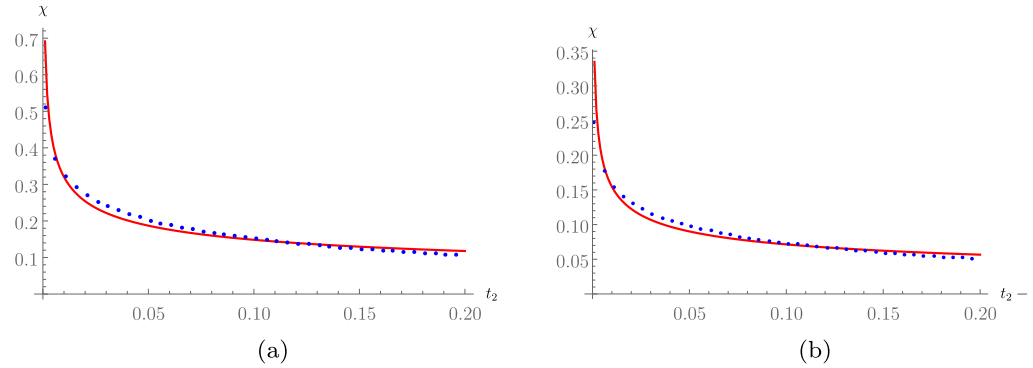


FIG. 6. The fidelity susceptibility χ_{MM} for the topological insulator, as a function of $(M - M_c)$. The points are the result of numerical integration. The red curve is the result of the fit. (a) The fidelity susceptibility χ_{MM} as a function of M in a neighborhood of $M \equiv t_2 = 0$. (b) The fidelity susceptibility χ_{MM} as a function of $(M - 2)$ in a neighborhood of $M \equiv t_2 = 2$.

V. SPECTRUM AND EDGE STATES FOR THE TOPOLOGICAL INSULATOR

The previous analysis from Secs. III and IV shows the absence of finite temperature topological phase transitions. In this section, we further analyze these results by studying the edge states, thus complementing our study.

Topological orders are characterized by Chern numbers, or similar global topological invariants that are associated with the homotopy class of the ground-state manifold of the Hamiltonian (having discrete values). Thus, they are inherently connected with the zero-temperature behavior of the system. To study the finite temperature behavior of the system, one has to search for other quantities and properties whose behavior at zero temperature coincide with that of the known topological invariants, and can be naturally generalized to finite temperatures. A feature associated with the topological order that can be studied at finite temperature are the edge states, whose behavior is described by their spatial occupation number.

Consider the Hamiltonian given by Eq. (14), defined on a 2D lattice, describing a topological insulator. We consider the x direction to be finite, with N sites, and with open boundary conditions. For the y direction we take periodic boundary conditions, thus allowing for a partial Fourier transformation along this direction. Topologically, the system is a cylinder; see Fig. 7.

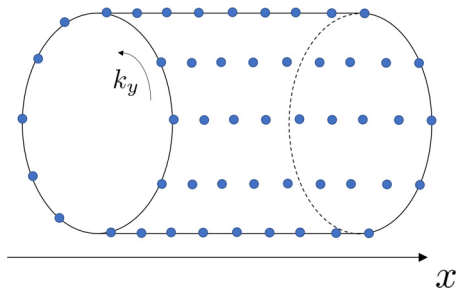


FIG. 7. To study the edge states, we consider the system on a cylinder topology. The quasimomentum winds around the periodic direction and the dots denote lattice sites.

The 2D lattice Hamiltonian then reduces to a family of 1D Hamiltonians parametrized by the quasimomentum k_y. More concretely, we take the Hamiltonian in Eq. (14) and we perform the Fourier transform in the j index, i.e., along the y axis. The fermion creation and annihilation operators in real space, described by the lattice coordinates (i, j), are then expanded in partial Fourier series,

$$c^\dagger_{i,j} = \frac{1}{\sqrt{N_y}} \sum_{k_y} \exp(+ik_y j) c^\dagger_i(k_y),$$

$$c_{i,j} = \frac{1}{\sqrt{N_y}} \sum_{k_y} \exp(-ik_y j) c_i(k_y),$$

where k_y is the quasimomentum along the y direction. The Hamiltonian can then be written in terms of the partial Fourier modes (see Appendix B),

$$\mathcal{H} = \sum_{k_y} \mathcal{H}(k_y) = \sum_{k_y} \sum_{i,j} c^\dagger_i(k_y) H_{i,j}(k_y) c_j(k_y), \qquad (24)$$

with $H_{i,j}(k_y)$ as given in Eq. (B8). For each k_y, the $N \times N$ matrix $H(k_y) = [H_{i,j}(k_y)]$ describes a 1D system. Now the energy eigenstates are classified into bulk states and edge states. The bulk states are delocalized, while the edge states are exponentially localized along the x direction, either on the left ($i = 1$) or on the right ($i = N$). If the bulk of the system is in a topologically nontrivial phase and we find energy eigenstates within the gap, by the bulk-to-boundary correspondence principle, they should be edge states canceling the anomaly coming from the bulk. Therefore, when plotting $E(k_y)$ for t_2 in a topologically nontrivial phase, we find zero-energy edge states (see Fig. 8 for $N = 20$). The effects of temperature on these states can be studied by considering the thermal states $\rho = \exp[-\beta(\mathcal{H} - \mu n)]/Z$, where β is the inverse of the temperature, n is the total number operator, and μ is the chemical potential. In the topologically trivial case the spectrum of the system decomposes into two sets of bands separated by a gap. In the zero-temperature limit and zero chemical potential, the lower bands, i.e., the bands below the Fermi level, are occupied and ρ becomes a projector onto this state.

In contrast, the spectrum of the topologically nontrivial system is composed of the previous two sets of bands and

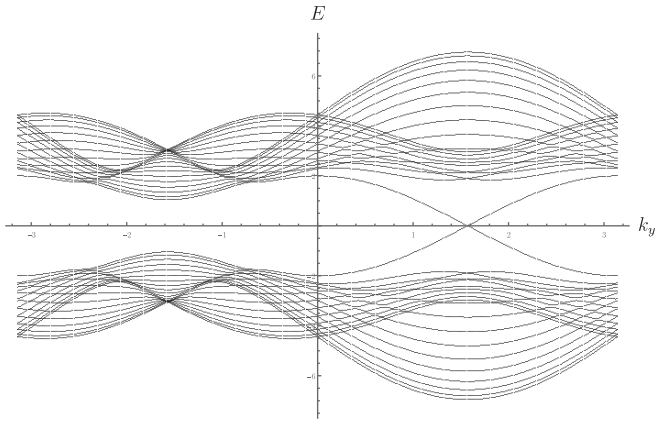


FIG. 8. Spectrum of the 2D TI, for $t_2 = 1.5$ in the cylinder topology. The size of the lattice in the x direction is 20 sites.

additional midgap states which are localized at the boundary of the system (cf. Fig. 8). By adding a small negative chemical potential μ (smaller than the gap), since the edge states are localized and are midgap states, the expectation value of the particle number $\langle n_i(k_y)\rangle$ on the zero-temperature limit of the density matrix will vary drastically at k_y near the critical momentum where the gap closes and i is at the edge, i.e., $i = 1$ and $i = N$. As the temperature raises, the particle numbers will saturate, since the density matrix tends to the totally mixed state. Indeed, this is what we can observe for the 2D TI in Fig. 9. A big variation in particle number occurs at the edges and in the vicinity of the critical momentum $k_y = \pi/2$, signaling the existence of the edge states.

Furthermore, our edge states analysis manifests the bulk-to-boundary principle. For the Chern insulator considered, the bulk-to-boundary principle states that the difference between the number of right and left chiral edge states in the phase boundary is equal to the difference in the Chern numbers between the phases. This results in having one right mover on the left boundary and one left mover on the right boundary, since we have $\Delta\mathrm{Ch} = \mp 1$. This can be clearly seen from Fig. 9.

To find the bulk spectrum we used Eq. (15). We can also write Eq. (15) as $H(k_x, k_y) = d^\mu(k_x, k_y)\sigma_\mu$, where $(d^\mu) = (d^x, d^y, d^z) = (2t_1\cos(k_x), 2t_1\cos(k_y), 2t_2\cos(k_x + k_y) + 2t_3[\sin(k_x) + \sin(k_y)])$ and $(\sigma_\mu) = (\sigma_x, \sigma_y, \sigma_z)$. The bulk spectrum consists of two bands given by $E(k_x, k_y) = \pm|\mathbf{d}|$. For fixed $t_1 = t_3 = 1$, the spectrum is plotted for various values of t_2, as shown in Fig. 10. We can see that the energy gap closes at the values $t_2 = -2,\ 0,\ +2$. This can be evaluated by noticing that the closing of the gap is equivalent to $\vec{d} = 0$. To have this, we need $d_x = 0$ and $d_y = 0$, which implies $\cos(k_x) = \cos(k_y) = 0$. By inspection of d_z, we find the following four different points in the Brillouin zone where gap closes. For $t_2 = -2$, the gap closes at $k_x = k_y = -\frac{\pi}{2}$, whereas for $t_2 = 0$, the gap closes at two different points $k_x = \frac{\pi}{2}, k_y = -\frac{\pi}{2}$ and $k_x = -\frac{\pi}{2}, k_y = \frac{\pi}{2}$ and for $t_2 = 2$ the gap closes at $k_x = k_y = \frac{\pi}{2}$ as shown in Fig. 10. For $t_2 \neq -2,\ 0,\ +2$ the system is gapped.

VI. CONCLUSIONS AND FUTURE WORK

We applied the fidelity and the Uhlmann connection analysis of phase transitions [18] to 2D free-fermion topological superconductors and insulators. First, we showed that the fidelity and the quantity Δ associated to the Uhlmann factor are detecting zero-temperature topological quantum phase transitions, as they assume nontrivial values around the gap-closing points. At finite temperatures, we probed both quantities with respect to small variations of the parameter of the Hamiltonian and the temperature. This twofold analysis confirmed the fact that the two quantities behave in a different way, with respect to changes of the spectrum and the eigenbasis. Specifically, when we kept the temperature constant, and varied the parameter of the Hamiltonian, which affects both the spectrum and the eigenbasis of the state, the fidelity and Δ were signaling the zero-temperature topological phase transitions. On the other hand, when we varied the temperature alone, which only affects the spectrum, the fidelity was still detecting the zero-temperature topological phase transitions, while Δ stayed trivial, since it is only sensitive to changes of the eigenbasis of the state.

Furthermore, we observed that as we increase the temperature, the nontriviality of fidelity and Δ around the gap-closing points at zero temperature is gradually smeared out and both quantities become trivial, in the infinite temperature limit. This indicates that there are no finite temperature phase transi-

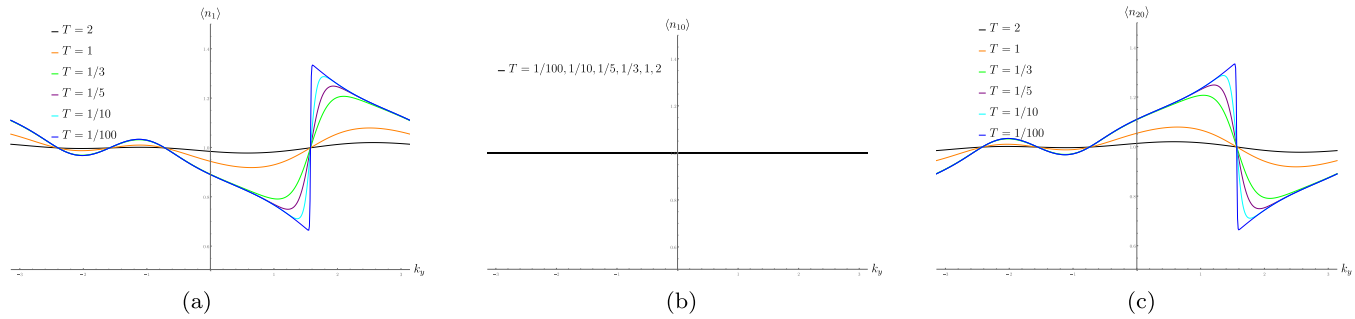
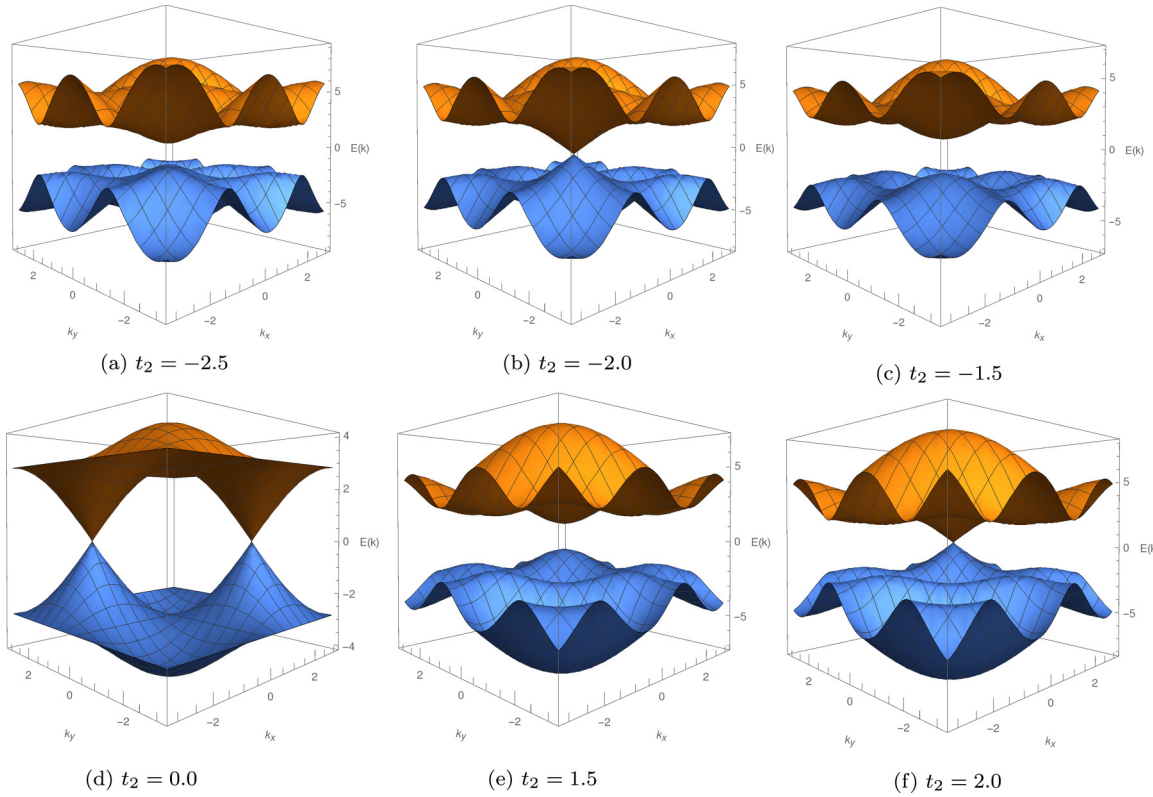


FIG. 9. Expectation value of the occupation number at the bulk and boundary as a function of the quasimomentum k_y for the 2D topological insulator. The results are obtained for a chain of 20 sites, $t_2 = 1.5$, a topologically nontrivial phase, $\mu = 0.01$ and temperatures $T = 1/100, 1/10, 1/5, 1/3, 1, 2$. (a) Occupation number at the left edge. (b) Occupation number in the bulk. (c) Occupation number at the right edge.



(a) $t_2 = -2.5$ (b) $t_2 = -2.0$ (c) $t_2 = -1.5$

(d) $t_2 = 0.0$ (e) $t_2 = 1.5$ (f) $t_2 = 2.0$

FIG. 10. The bulk dispersion relation at different values of t_2 for the 2D topological insulator. In (a), (c), and (e), we show the cases of the system with a bulk gap (insulating phase), while in (b), (d), and (f) the system has no bulk gap.

tions and the topological features of the systems at $T = 0$ are gradually washed away. Furthermore, we analyzed the noncommutativity of the thermodynamic and zero-temperature limits of the fidelity susceptibility and gave an analytical proof of the absence of finite temperature phase transitions in the class of models considered. This proof also complements the results presented in our previous work of Ref. [18]. Moreover, we analyzed the critical behavior at zero temperature by looking at the critical exponents of the fidelity susceptibility in the vicinity of the critical points.

We also performed a detailed study of the edge states of the systems at finite temperature. Our results show that as the temperature is increased, the $T = 0$ prominent edge states start mixing with the bulk states, confirming the conclusion of the fidelity and Δ analysis, i.e., the gradual smearing of the topological features at finite temperatures (see also a related recent study [37]).

Finally, we would like to point out some directions of future work. First, one may consider the case of Hamiltonians with interactions, or systems of bosons. A slightly different direction would be to use the same approach in the case of open systems, in which also the system-bath interaction should be taken into account. Moreover, the chiral p-wave superconductor that we considered in our analysis is shown to host vortices that obey non-Abelian statistics, which are the main constituent of proposals for fault-tolerant quantum computation [38,39] and the design of stable quantum memories [40], due to their robustness against perturbations of the Hamiltonian. While the robustness of these modes has already been considered in the literature, it would be relevant to perform a more detailed quantitative analysis of their behavior with respect to temperature, using our approach for the study of the edge states. Additionally, in this case the zero energy Majorana bound states appearing due to the vortices should be relevant in the behavior of the Uhlmann factor and deserve a separate and thorough study.

ACKNOWLEDGMENTS

S.T.A., B.M., and C.V. acknowledge the support from DP-PMI and FCT (Portugal) through Grants No. PD/BD/113651/2015, No. SFRH/BD/52244/2013, and No. PD/BD/52652/2014, respectively. B.M. acknowledges the support from Fundação para a Ciência e Tecnologia (Portugal) namely through PTDC/POPH/POCH programs and projects UID/EEA/50008/2013, IT/QuSim, IT/QuNet, ProQuNet, partially funded by EU FEDER, from the EU FP7 project PAPETS (GA 323901) and from the JTF project NQuN (ID 60478). N.P. acknowledges the IT project QbigD funded by FCT PEst-OE/EEI/LA0008/2013, UID/EEA/50008/2013, and the Confident project PTDC/EEI-CTP/4503/2014. Support from FCT (Portugal) through Grant No. UID/CTM/04540/2013 and the support of SQIG – Security and Quantum Information Group, are also acknowledged.

APPENDIX A: ANALYTIC EXPRESSION FOR FIDELITY

The fidelity between two thermal states of quasifree fermions,

$$\rho = \exp\left(-\tfrac{1}{2}\beta\widetilde{\psi}^{\dagger}H_{\mathrm{BdG}}\widetilde{\psi}\right)/Z, \tag{A1}$$

where $\widetilde{\psi} = (\psi_1, \ldots, \psi_N, \psi_1^\dagger, \ldots, \psi_N^\dagger)^T$, N is the total number of degrees of freedom on the system, and H_{BdG} has the BdG form, is given by

$$F(\rho, \rho') = \mathrm{Tr}\Big(\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\Big) = \sqrt{\frac{\det(I + e^{C_{\mathrm{BdG}}/2})}{\det^{1/2}(I + e^{-\beta H_{\mathrm{BdG}}})\det^{1/2}(I + e^{-\beta H'_{\mathrm{BdG}}})}}, \quad \text{(A2)}$$

where C_{BdG} is defined by

$$e^{-\beta H_{\mathrm{BdG}}/2} e^{-\beta H'_{\mathrm{BdG}}} e^{-\beta H_{\mathrm{BdG}}/2} = e^{C_{\mathrm{BdG}}}. \quad \text{(A3)}$$

The derivation of this result proceeds as follows. By writing Majorana operators

$$\gamma_{2i-1} = \psi_i + \psi_i^\dagger, \qquad \gamma_{2i} = (1/i)(\psi_i - \psi_i^\dagger),\ i = 1, \ldots, N,$$

which satisfy the Clifford algebra relations

$$\{\gamma_i, \gamma_j\} = 2\delta_{ij}, \quad \text{for } i, j = 1, \ldots, N,$$

one can show that

$$\mathcal{H} = \frac{i}{4}\gamma A \gamma^T,$$

where $\gamma = (\gamma_1, \ldots, \gamma_{2N})$, $iA = R H_{\mathrm{BdG}} R^{-1}$ for an invertible matrix R, and $A \in \mathfrak{so}(2N)$, i.e., A is real and skew symmetric. The matrix A can be skew diagonalized,

$$A = S\Big[\oplus_{i=1}^N \begin{pmatrix} 0 & \varepsilon_i \\ -\varepsilon_i & 0 \end{pmatrix}\Big] S^{-1},$$

with $\varepsilon_i \in \mathbb{R}$, $i = 1, \ldots, N$, and $S \in \mathrm{SO}(2N)$. The rotated $\gamma' = \gamma S$ satisfy the same Clifford algebra relations, and one can define new canonical fermion operators,

$$\gamma'_{2i-1} = \psi'_i + \psi_i'^\dagger, \quad \gamma'_{2i} = (1/i)(\psi'_i - \psi_i'^\dagger),\ i = 1, \ldots, N,$$

so that

$$\mathcal{H} = \frac{i}{4}\sum_{i=1}^N \varepsilon_i(\gamma'_{2i-1}\gamma'_{2i} - \gamma'_{2i}\gamma'_{2i-1}) = \frac{i}{2}\sum_{i=1}^N \varepsilon_i \gamma'_{2i-1}\gamma'_{2i} = \frac{1}{2}\sum_{i=1}^N \varepsilon_i(\psi_i'^\dagger \psi'_i - \psi'_i \psi_i'^\dagger) = \sum_{i=1}^N \varepsilon_i \psi_i'^\dagger \psi'_i - \frac{1}{2}\sum_{i=1}^N \varepsilon_i.$$

It is now clear that

$$\mathrm{Tr}(e^{-\beta\mathcal{H}}) = \prod_{i=1}^N (e^{-\beta\varepsilon_i/2} + e^{\beta\varepsilon_i/2}).$$

Since iA and H are related by conjugation by R, their eigenvalues are the same, i.e., $\{\pm\varepsilon_i\}_{i=1}^N$. Therefore,

$$\det(e^{-\beta H_{\mathrm{BdG}}/2} + e^{\beta H_{\mathrm{BdG}}/2}) = \left(\prod_{i=1}^N (e^{-\beta\varepsilon_i/2} + e^{\beta\varepsilon_i/2})\right)^2,$$

and we then have

$$\mathrm{Tr}(e^{-\beta\mathcal{H}}) = \mathrm{Tr}(e^{-(\beta/2)\widetilde{\psi}^\dagger H_{\mathrm{BdG}}\widetilde{\psi}}) = \det^{1/2}(e^{-\beta H_{\mathrm{BdG}}/2} + e^{\beta H_{\mathrm{BdG}}/2}).$$

We need an additional ingredient. Consider A a matrix of the BdG type. Then, one can show, by using the commutation relations and $PAP = -A^T$, with P the particle-hole conjugation matrix, so $\widetilde{\psi}^\dagger P = \widetilde{\psi}^T$, that the following relations hold:

$$e^{(1/2)\widetilde{\psi}^\dagger A\widetilde{\psi}}\widetilde{\psi}^\dagger e^{-(1/2)\widetilde{\psi}^\dagger A\widetilde{\psi}} = \widetilde{\psi}^\dagger e^A,$$
$$e^{(1/2)\widetilde{\psi}^\dagger A\widetilde{\psi}}\widetilde{\psi} e^{-(1/2)\widetilde{\psi}^\dagger A\widetilde{\psi}} = e^{-A}\widetilde{\psi}.$$

As a consequence, if A and B are of the BdG type,

$$e^{(1/2)\widetilde{\psi}^\dagger A\widetilde{\psi}} e^{(1/2)\widetilde{\psi}^\dagger B\widetilde{\psi}}\widetilde{\psi}^\dagger e^{-(1/2)\widetilde{\psi}^\dagger B\widetilde{\psi}} e^{-(1/2)\widetilde{\psi}^\dagger A\widetilde{\psi}} = \widetilde{\psi}^\dagger e^A e^B,$$

and similarly for $\widetilde{\psi}$. Since we are working over the field of complex numbers, we can find a C_{BdG} such that

$$e^{-\beta(H_{\mathrm{BdG}}/2)} e^{-\beta H'_{\mathrm{BdG}}} e^{-\beta(H_{\mathrm{BdG}}/2)} = e^{C_{\mathrm{BdG}}}.$$

Moreover, this C_{BdG} can be chosen to satisfy $PC_{\mathrm{BdG}}^T P = -C_{\mathrm{BdG}}$. This can be seen from

$$P(e^A e^B)^T P = e^{PB^T P} e^{PA^T P} = (e^A e^B)^{-1} = e^{PC^T P}.$$

It then follows that

$$e^{(1/2)\widetilde{\psi}^\dagger A\widetilde{\psi}} e^{(1/2)\widetilde{\psi}^\dagger B\widetilde{\psi}} = e^{(1/2)\widetilde{\psi}^\dagger C\widetilde{\psi}},$$

since the action of $e^{(1/2)\widetilde{\psi}^\dagger A\widetilde{\psi}} e^{(1/2)\widetilde{\psi}^\dagger B\widetilde{\psi}}$ is completely determined by its action on $\widetilde{\psi}^\dagger$.

Now, we can compute the fidelity. It reads

$$F(\rho, \rho') = \mathrm{Tr}(\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}) = \mathrm{Tr}\left(\sqrt{\frac{e^{-(\beta/2)\widetilde{\psi}^\dagger(H_{\mathrm{BdG}}/2)\widetilde{\psi}} e^{-(\beta/2)\widetilde{\psi}^\dagger H'_{\mathrm{BdG}}\widetilde{\psi}} e^{-(\beta/2)\widetilde{\psi}^\dagger(H_{\mathrm{BdG}}/2)\widetilde{\psi}}}{ZZ'}}\right).$$

We have, because of the previous results,

$$\begin{aligned} F(\rho,\rho') &= \mathrm{Tr}\left(\sqrt{\frac{e^{-(1/2)\widetilde{\psi}^\dagger C_{\mathrm{BdG}}\widetilde{\psi}}}{ZZ'}}\right) \\ &= \mathrm{Tr}\left(\frac{e^{-(1/2)\widetilde{\psi}^\dagger (C_{\mathrm{BdG}}/2)\widetilde{\psi}}}{\sqrt{ZZ'}}\right) \\ &= \frac{\det^{1/2}(e^{C_{\mathrm{BdG}}/4}+e^{-C_{\mathrm{BdG}}/4})}{\det^{1/4}(e^{\beta H_{\mathrm{BdG}}/2}+e^{-\beta H_{\mathrm{BdG}}/2})\det^{1/4}(e^{\beta H'_{\mathrm{BdG}}/2}+e^{-\beta H'_{\mathrm{BdG}}/2})}. \end{aligned} \tag{A4}$$

One can then write

$$F(\rho,\rho') = \frac{\det^{1/2}(e^{-C_{\mathrm{BdG}}/4})\det^{1/2}(I+e^{C_{\mathrm{BdG}}/2})}{\det^{1/4}(e^{\beta H_{\mathrm{BdG}}/2}e^{\beta H'_{\mathrm{BdG}}/2})\det^{1/4}(I+e^{-\beta H_{\mathrm{BdG}}})\det^{1/4}(I+e^{-\beta H'_{\mathrm{BdG}}})}.$$

But,

$$\det^{1/2}(e^{-C_{\mathrm{BdG}}/4}) = \det^{1/8}(e^{-C_{\mathrm{BdG}}}) \text{ and}$$
$$\det^{1/2}(e^{\beta H_{\mathrm{BdG}}/2}e^{\beta H'_{\mathrm{BdG}}/2}) = \det^{1/4}(e^{\beta H_{\mathrm{BdG}}/4}e^{\beta H'_{\mathrm{BdG}}/2}e^{\beta H_{\mathrm{BdG}}/4}) = \det^{1/4}(e^{-C_{\mathrm{BdG}}/2}) = \det^{1/8}(e^{-C_{\mathrm{BdG}}}).$$

Therefore,

$$F(\rho,\rho') = \sqrt{\frac{\det(I+e^{C_{\mathrm{BdG}}/2})}{\det^{1/2}(I+e^{-\beta H_{\mathrm{BdG}}})\det^{1/2}(I+e^{-\beta H'_{\mathrm{BdG}}})}}.$$

In the main text, we consider the fidelity for topological insulators and superconductors in translation invariant systems. Let us denote the fields $\widetilde{\psi}_i(\mathbf{x})$, where i labels internal degrees of freedom and $\mathbf{x}$ the position. Since we have translation invariance, we can go into momentum space,

$$\widetilde{\psi}_i(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\widetilde{\psi}_i(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}},$$

where V denotes the number of lattice sites. However, since $\widetilde{\psi}_i(\mathbf{x}) = (\psi_i(\mathbf{x}), \psi_i^\dagger(\mathbf{x}))^T$, where the second component is the adjoint of the first, it follows that $\widetilde{\psi}_i(\mathbf{k}) = (\psi_i(\mathbf{k}), \psi_i^\dagger(-\mathbf{k}))^T$. Moreover, it is easy to see that the Fourier components of $H_{\mathrm{BdG}} = [H_{\mathrm{BdG}}(\mathbf{x},\mathbf{x}')]$ are of the form

$$H_{\mathrm{BdG}}(\mathbf{k}) = \begin{pmatrix} H(\mathbf{k}) & \Delta(\mathbf{k}) \\ \Delta^\dagger(\mathbf{k}) & -H^T(-\mathbf{k}) \end{pmatrix},$$

where $H(\mathbf{k})$ and $\Delta(\mathbf{k})$ are the Fourier components of H and Δ, respectively. It is now clear that in this setting

$$F(\rho,\rho') = \prod_{\mathbf{k}}\sqrt{\frac{\det(I+e^{C_{\mathrm{BdG}}(\mathbf{k})/2})}{\det^{1/2}(I+e^{-\beta H_{\mathrm{BdG}}(\mathbf{k})})\det^{1/2}(I+e^{-\beta H'_{\mathrm{BdG}}(\mathbf{k})})}},$$

where $C_{\mathrm{BdG}}(\mathbf{k})$ are the Fourier components of C_{BdG} (which is translation invariant). Moreover, in the charge conserving case, for the topological insulator $\Delta \equiv 0$ and

$$H_{\mathrm{BdG}}(\mathbf{k}) = \begin{pmatrix} H(\mathbf{k}) & 0 \\ 0 & -H^T(-\mathbf{k}) \end{pmatrix}.$$

The same form is valid for $C_{\mathrm{BdG}}(\mathbf{k})$, since the exponential of a block diagonal matrix is also block diagonal.

In this case, going back to Eq. (A4), we have

$$F(\rho,\rho') = \prod_{\mathbf{k}}\frac{\det^{1/2}(e^{C_{\mathrm{BdG}}(\mathbf{k})/4}+e^{-C_{\mathrm{BdG}}(\mathbf{k})/4})}{\det^{1/4}(e^{\beta H_{\mathrm{BdG}}(\mathbf{k})/2}+e^{-\beta H_{\mathrm{BdG}}(\mathbf{k})/2})\det^{1/4}(e^{\beta H'_{\mathrm{BdG}}(\mathbf{k})/2}+e^{-\beta H'_{\mathrm{BdG}}(\mathbf{k})/2})}.$$

Notice that

$$\begin{aligned} \prod_{\mathbf{k}}\det^{1/4}(e^{\beta H_{\mathrm{BdG}}(\mathbf{k})/2}+e^{-\beta H_{\mathrm{BdG}}(\mathbf{k})/2}) &= \prod_{\mathbf{k}}\det^{1/4}(e^{\beta H(\mathbf{k})/2}+e^{-\beta H(\mathbf{k})/2})\det^{1/4}\left(e^{-\beta H^T(-\mathbf{k})/2}+e^{\beta H^T(-\mathbf{k})/2}\right) \\ &= \prod_{\mathbf{k}}\det^{1/2}(e^{\beta H(\mathbf{k})/2}+e^{-\beta H(\mathbf{k})/2}), \end{aligned}$$

where we used invariance of the determinant by transposition. The same type of formulas hold for H'_{BdG} and C_{BdG}. Finally, we get the result

$$F(\rho,\rho') = \prod_{\mathbf{k}} \frac{\det(I + e^{C(\mathbf{k})/2})}{\det^{1/2}(I + e^{-\beta H(\mathbf{k})})\det^{1/2}(I + e^{-\beta H'(\mathbf{k})})}.$$

In terms of the vector $d^{\mu}(\mathbf{k};M)$, appearing in the topological insulator and in the topological superconductor models in the main text, we can write the fidelity as follows. Write a matrix $C(\mathbf{k})$ satisfying

$$e^{-(\beta/2)d^{\mu}(\mathbf{k};M)\sigma_{\mu}} e^{-\beta d^{\mu}(\mathbf{k};M')\sigma_{\mu}} e^{-(\beta/2)d^{\mu}(\mathbf{k};M)\sigma_{\mu}} = e^{C(\mathbf{k})}.$$

The explicit form of $C(\mathbf{k})$ is given in the Supplemental Material of [18]. Then, for the topological insulator,

$$F(\rho,\rho') = \prod_{\mathbf{k}} \frac{\det(I + e^{C(\mathbf{k})/2})}{\det^{1/2}\left(I + e^{-\beta d^{\mu}(\mathbf{k};M)\sigma_{\mu}}\right)\det^{1/2}\left(I + e^{-\beta d^{\mu}(\mathbf{k};M')\sigma_{\mu}}\right)}.$$

For the topological superconductor, since $\mathcal{H} = (1/2)\sum_{\mathbf{k}} \widetilde{\psi}^{\dagger}(\mathbf{k}) d^{\mu}(\mathbf{k};M)\sigma_{\mu}\widetilde{\psi}(\mathbf{k})$, we have

$$F(\rho,\rho') = \prod_{\mathbf{k}} \sqrt{\frac{\det(I + e^{C(\mathbf{k})/2})}{\det^{1/2}\left(I + e^{-\beta d^{\mu}(\mathbf{k};M)\sigma_{\mu}}\right)\det^{1/2}\left(I + e^{-\beta d^{\mu}(\mathbf{k};M')\sigma_{\mu}}\right)}}.$$

In other words, the two mathematical expressions differ by a square root. The reason for this lies in the fact that, while in the case of the topological superconductor, the Pauli matrices are acting on the particle-hole degrees of freedom, in the topological insulator case, the Pauli matrices act on the additional pseudospin degree of freedom.

APPENDIX B: HAMILTONIAN TRANSFORMATIONS

Let $A = t_1\sigma_x + it_3\sigma_z$, $B = t_1\sigma_y + it_3\sigma_z$ and $C = t_2\sigma_z$. Then,

$$\sum_j c^{\dagger}_{i+1,j} A c_{i,j} = \sum_{k_y} c^{\dagger}_{i+1,k_y} A c_{i,k_y}, \tag{B1}$$

$$\sum_j c^{\dagger}_{i,j+1} B c_{i,j} = \sum_{k_y} \exp(ik_y) c^{\dagger}_{i,k_y} B c_{i,k_y}, \tag{B2}$$

$$\sum_j c^{\dagger}_{i+1,j+1} C c_{i,j} = \sum_{k_y} \exp(ik_y) c^{\dagger}_{i+1,k_y} C c_{i,k_y}. \tag{B3}$$

By taking the Hermitian conjugate of the above three equations, we get

$$\sum_j c^{\dagger}_{i,j} A^{\dagger} c_{i+1,j} = \sum_{k_y} c^{\dagger}_{i,k_y} A^{\dagger} c_{i+1,k_y}, \tag{B4}$$

$$\sum_j c^{\dagger}_{i,j} B^{\dagger} c_{i,j+1} = \sum_{k_y} \exp(-ik_y) c^{\dagger}_{i,k_y} B^{\dagger} c_{i,k_y}, \tag{B5}$$

$$\sum_j c^{\dagger}_{i,j} C^{\dagger} c_{i+1,j+1} = \sum_{k_y} \exp(-ik_y) c^{\dagger}_{i,k_y} C^{\dagger} c_{i+1,k_y}. \tag{B6}$$

Summing Eqs. (B1) and (B3), we get

$$\sum_{k_y} c^{\dagger}_{i+1,k_y} A c_{i,k_y} + \sum_{k_y} \exp(ik_y) c^{\dagger}_{i+1,k_y} C c_{i,k_y} = \sum_{k_y} c^{\dagger}_{i+1,k_y} [t_1\sigma_x + it_3\sigma_z + t_2\exp(ik_y)\sigma_z] c_{i,k_y}.$$

Accordingly, from Eqs. (B2) and (B5)

$$\sum_{k_y} \exp(ik_y) c^{\dagger}_{i,k_y} B c_{i,k_y} + \sum_{k_y} \exp(-ik_y) c^{\dagger}_{i,k_y} B^{\dagger} c_{i,k_y} = \sum_{k_y} c^{\dagger}_{i,k_y} [2\cos(k_y) t_1\sigma_y - 2\sin(k_y) t_3\sigma_z] c_{i,k_y},$$

and finally from Eqs. (B4) and (B6)

$$\sum_{k_y} c^{\dagger}_{i,k_y} A^{\dagger} c_{i+1,k_y} + \sum_{k_y} \exp(-ik_y) c^{\dagger}_{i,k_y} C^{\dagger} c_{i+1,k_y} = \sum_{k_y} c^{\dagger}_{i,k_y} [t_1\sigma_x - it_3\sigma_z + t_2\exp(-ik_y)\sigma_z] c_{i+1,k_y}.$$

Thus the total Hamiltonian of the topological insulator with transitional invariance in the y axis becomes

$$\begin{aligned} H = \sum_{k_y}\sum_i \big\{ & c^{\dagger}_{i+1,k_y}[t_1\sigma_x + it_3\sigma_z + t_2\exp(ik_y)\sigma_z] c_{i,k_y} \\ & + c^{\dagger}_{i,k_y}[2\cos(k_y) t_1\sigma_y - 2\sin(k_y) t_3\sigma_z] c_{i,k_y} \\ & + c^{\dagger}_{i,k_y}[t_1\sigma_x - it_3\sigma_z + t_2\exp(-ik_y)\sigma_z] c_{i+1,k_y} \big\}, \end{aligned}$$

and for $H(k_y)$ we have

$$\begin{aligned} H(k_y) = \sum_{i,j} \big\{ & c^{\dagger}_{i,k_y}[t_1\sigma_x + it_3\sigma_z + t_2\exp(ik_y)\sigma_z]\delta_{i,j+1} c_{j,k_y} \\ & + c^{\dagger}_{i,k_y}[2\cos(k_y) t_1\sigma_y - 2\sin(k_y) t_3\sigma_z]\delta_{i,j} c_{j,k_y} \\ & + c^{\dagger}_{i,k_y}[t_1\sigma_x - it_3\sigma_z + t_2\exp(-ik_y)\sigma_z]\delta_{i+1,j} c_{j,k_y} \big\}, \end{aligned}$$

with the (i,j) matrix elements of $H(k_y)$ being

$$\begin{aligned} H_{i,j}(k_y) = {} & [t_1\sigma_x + it_3\sigma_z + t_2\exp(ik_y)\sigma_z]\delta_{i,j+1} \\ & + [2\cos(k_y) t_1\sigma_y - 2\sin(k_y) t_3\sigma_z]\delta_{i,j} \\ & + [t_1\sigma_x - it_3\sigma_z + t_2\exp(-ik_y)\sigma_z]\delta_{i+1,j}. \end{aligned} \tag{B7}$$

Simplifying Eq. (B7), we obtain

$$\begin{aligned} H_{i,j}(k_y) = {} & t_1(\delta_{i,j+1} + \delta_{i+1,j})\sigma_x + 2t_1\cos(k_y)\delta_{i,j}\sigma_y \\ & + \{i[t_3 + t_2\sin(k_y)](\delta_{i,j+1} - \delta_{i+1,j}) \\ & - 2t_3\sin(k_y)\delta_{i,j} + t_2\cos(k_y)(\delta_{i,j+1} + \delta_{i+1,j})\}\sigma_z, \end{aligned} \tag{B8}$$

or

$$H_{i,j}(k_y) = a\sigma_x + b\sigma_y + c\sigma_z, \tag{B9}$$

with $a = t_1(\delta_{i,j+1} + \delta_{i+1,j})$, $b = 2t_1\cos(k_y)\delta_{i,j}$, and $c = i[t_3 + t_2\sin(k_y)](\delta_{i,j+1} - \delta_{i+1,j}) - 2t_3\sin(k_y)\delta_{i,j} + t_2\cos(k_y)(\delta_{i,j+1} + \delta_{i+1,j})$.

If we also consider periodic boundary conditions along the x axis, we will have translational invariance, and after applying the full Fourier transform, we get the following bulk Hamiltonian for the topological insulator in two dimensions:

$$H(\mathbf{k}) = 2t_1\cos(k_x)\sigma_x + 2t_1\cos(k_y)\sigma_y + \{2t_2\cos(k_x + k_y) + 2t_3[\sin(k_x) + \sin(k_y)]\}\sigma_z.$$

APPENDIX C: DERIVATION OF THE BURES METRIC

Consider a generic charge symmetric fermionic Gaussian state

$$\rho = \exp(X), \text{ with } X = -\psi^\dagger H\psi - \ln Z,\ \ln Z = \mathrm{Tr}e^{-\psi^\dagger H\psi}, \text{ and } H \in \mathrm{Herm}(\mathbb{C}^N).$$

Here ψ and $\psi^\dagger$ are the usual arrays of fermion destruction and creation operators. We are interested in evaluating the pullback of the Bures metric by the map $\rho : \mathrm{Herm}(\mathbb{C}^N) \ni H \mapsto e^X$. Now, the Bures metric is given by

$$ds^2 = \mathrm{Tr}\{\rho G^2\},$$

with

$$d\rho = G\rho + \rho G.$$

The last equation is equivalent to (for invertible states, which is the case at hand)

$$\rho^{-1}d\rho = \rho^{-1}G\rho + G = (1 + e^{-\mathrm{ad}_X})(G),$$

where $\mathrm{ad}_X(\cdot) := [X, \cdot]$. Now,

$$d\rho = de^X = e^X(e^{-X}de^X) = e^X\int_0^1 ds\, e^{-sX}dXe^{sX},$$

since

$$\frac{\partial}{\partial s}[e^{-sX}d(e^{sX})] = e^{-sX}dXe^{sX}.$$

Therefore,

$$d\rho = e^X\int_0^1 ds\, e^{-s\,\mathrm{ad}_X}(dX) = -e^X(\mathrm{ad}_X)^{-1}(e^{-\mathrm{ad}_X} - 1)(dX),$$

and

$$\rho^{-1}d\rho = (\mathrm{ad}_X)^{-1}(1 - e^{-\mathrm{ad}_X})(dX).$$

Thus,

$$(1 + e^{-\mathrm{ad}_X})(G) = (\mathrm{ad}_X)^{-1}(1 - e^{-\mathrm{ad}_X})(dX)$$

Formally,

$$G = \frac{1}{2}\frac{\tanh\left(\frac{\mathrm{ad}_X}{2}\right)}{\frac{\mathrm{ad}_X}{2}}(dX).$$

Notice that

$$dX = -\psi^\dagger dH\psi - \frac{dZ}{Z}.$$

Using the fact that any Hermitian matrix can be diagonalized by a unitary matrix, we can write

$$H = SDS^{-1}, \text{ with } S \in \mathrm{U}(N) \text{ and } D = \mathrm{diag}(\lambda_1, \ldots, \lambda_N).$$

then,

$$dH = S([S^{-1}dS, D] + dD)S^{-1}.$$

Then,

$$\psi^\dagger S = \widetilde{\psi}^\dagger \text{ and } S^{-1}\psi = \widetilde{\psi},$$

satisfy the usual canonical anticommutation relations. Moreover,

$$\mathrm{ad}_X(\widetilde{\psi}_i^\dagger) = -\lambda_i\widetilde{\psi}_i^\dagger,\ i = 1, \ldots, N,$$
$$\mathrm{ad}_X(\widetilde{\psi}_i) = \lambda_i\widetilde{\psi}_i,\ i = 1, \ldots, N.$$

We can then write

$$dX = -\widetilde{\psi}^\dagger([S^{-1}dS, D] + dD)\widetilde{\psi} - \frac{dZ}{Z}.$$

and

$$G = \frac{1}{2}\frac{\tanh\left(\frac{\mathrm{ad}_X}{2}\right)}{\frac{\mathrm{ad}_X}{2}}(dX)$$
$$= -\frac{1}{2}\sum_{i,j=1}^N \widetilde{\psi}_i^\dagger\left(-(S^{-1}dS)_{ij}\frac{\tanh\left(\frac{\lambda_i-\lambda_j}{2}\right)}{\frac{\lambda_i-\lambda_j}{2}}(\lambda_i - \lambda_j) + d\lambda_i\delta_{ij}\right)\widetilde{\psi}_j - \frac{1}{2}\frac{dZ}{Z}$$
$$= \sum_{i,j=1}^N \widetilde{\psi}_i^\dagger\left[(S^{-1}dS)_{ij}\tanh\left(\frac{\lambda_i - \lambda_j}{2}\right) - \frac{d\lambda_i}{2}\delta_{ij}\right]\widetilde{\psi}_j - \frac{1}{2}\frac{dZ}{Z},$$

so, clearly, G is of the form

$$G = \psi^\dagger K\psi + C,$$

for a Hermitian matrix valued one-form $K = SBS^{-1}$ and one-form C,

$$B_{ij} = (S^{-1}dS)_{ij} \tanh\left(\frac{\lambda_i - \lambda_j}{2}\right) - \frac{d\lambda_i}{2}\delta_{ij},$$
$$C = -\frac{1}{2}\frac{dZ}{Z}.$$

The one-form $S^{-1}dS$ is the Maurer-Cartan differential form. To evaluate the metric tensor, we just need to compute the trace,

$$ds^2 = \mathrm{Tr}(\rho G^2) = \langle G^2\rangle = \langle\psi^\dagger K\psi\psi^\dagger K\psi\rangle + 2\langle\psi^\dagger K\psi\rangle C + C^2.$$

To evaluate the last expression, consider the correlation function

$$g_{ji} := \langle\psi_i^\dagger\psi_j\rangle.$$

It is clear that for any single-particle observable $A \in \mathrm{Herm}(\mathbb{C}^n)$, we have

$$\langle\psi^\dagger A\psi\rangle = \mathrm{Tr}(Ag).$$

Moreover, we will use Wick's theorem to evaluate expressions of the form

$$\langle\psi_i^\dagger\psi_j\psi_k^\dagger\psi_l\rangle = g_{ji}g_{lk} + g_{li}(\delta_{jk} - g_{jk}).$$

Then, we can write

$$\begin{aligned}\langle\psi^\dagger K\psi\psi^\dagger K\psi\rangle &= K_{ij}K_{kl}\langle\psi_i^\dagger\psi_j\psi_k^\dagger\psi_l\rangle = K_{ij}K_{kl}(g_{ji}g_{lk} + g_{li}(\delta_{jk} - g_{jk})\\ &= [\mathrm{Tr}(Kg)]^2 + \mathrm{Tr}(K^2g) - \mathrm{Tr}(gKgK),\end{aligned}$$

and

$$\langle\psi^\dagger K\psi\rangle = \mathrm{Tr}(Kg).$$

Moreover, since $\mathrm{Tr}(d\rho) = 0$, we have

$$0 = 2\mathrm{Tr}(\rho G) = 2(\langle\psi^\dagger K\psi\rangle + C),$$

hence

$$C = -\langle\psi^\dagger K\psi\rangle = -\mathrm{Tr}(Kg) = -\frac{1}{2}\frac{dZ}{Z}.$$

Finally, we can write

$$\begin{aligned}ds^2 &= [\mathrm{Tr}(Kg)]^2 + \mathrm{Tr}(K^2g) - \mathrm{Tr}(gKgK) - 2[\mathrm{Tr}(Kg)]^2 + [\mathrm{Tr}(Kg)]^2\\ &= \mathrm{Tr}(K^2g) - \mathrm{Tr}(gKgK).\end{aligned}$$

Also as

$$ds^2 = \langle\psi^\dagger K\psi\psi^\dagger K\psi\rangle + 2\langle\psi^\dagger K\psi\rangle C + C^2 = \langle(\psi^\dagger K\psi)^2\rangle - \langle\psi^\dagger K\psi\rangle^2.$$

Next, notice that

$$\langle\widetilde{\psi}_j^\dagger\widetilde{\psi}_i\rangle = \delta_{ij}\frac{1}{e^{-\lambda_i}+1} \equiv \delta_{ij}n_i = S_{kj}\overline{S}_{li}\langle\psi_k^\dagger\psi_l\rangle = S_{il}^{-1}g_{lk}S_{kj},$$

and hence,

$$g = (1 + e^{-H})^{-1}.$$

Now, defining $A_{ij} = (S^{-1}dS)_{ij} = -\overline{A}_{ji}$ and $t_{ij} = -t_{ji} = \tanh[(\lambda_i - \lambda_j)/2]$, we have

$$\begin{aligned}\mathrm{Tr}(K^2g) = B_{ij}B_{jl}n_l\delta_{li} = n_iB_{ij}B_{ji} &= n_i\Big[A_{ij}t_{ij} - \frac{d\lambda_i}{2}\delta_{ij}\Big]\Big[A_{ji}t_{ji} - \frac{d\lambda_j}{2}\delta_{ji}\Big]\\ &= n_i\left(|A_{ij}|^2t_{ij}^2 + \frac{1}{4}d\lambda_i^2\right),\end{aligned}$$

and

$$\begin{aligned}\mathrm{Tr}\big(gKgK\big) = n_i\delta_{ij}B_{jk}n_k\delta_{kl}B_{li} &= n_in_jB_{ij}B_{ji}\\ &= n_in_j\left[A_{ij}t_{ij} - \frac{d\lambda_i}{2}\delta_{ij}\right]\left[A_{ji}t_{ji} - \frac{d\lambda_j}{2}\delta_{ji}\right]\\ &= n_in_j\left(|A_{ij}|^2t_{ij}^2 + \frac{1}{4}d\lambda_i^2\delta_{ij}\right),\end{aligned}$$

hence,

$$ds^2 = \sum_{i\neq j} n_i(1-n_j)|A_{ij}|^2 t_{ij}^2 + \frac{1}{4}\sum_{i=1}^{N} n_i(1-n_i)d\lambda_i^2$$

$$= \sum_{i\neq j} n_i(1-n_j)|(S^{-1}dS)_{ij}|^2 \tanh^2\left(\frac{\lambda_i-\lambda_j}{2}\right) + \frac{1}{4}\sum_{i=1}^{N} n_i(1-n_i)d\lambda_i^2.$$

The second term is the classical one, i.e., the classical Fisher information associated with the probability distribution,

$$p_I := p_{i_1,\cdots,i_p} = \frac{e^{-(\lambda_{i_1}+\cdots+\lambda_{i_p})}}{Z},\ 1<i_1<\cdots<i_p<N.$$

Since

$$dp_I = -\big[\big(d\lambda_{i_1}+\cdots+d\lambda_{i_p}\big)+d\ln(Z)\big]p_I,$$

we have

$$\frac{1}{4}\sum_I \frac{dp_I^2}{p_I} = \frac{1}{4}\sum_I p_I\big[\big(d\lambda_{i_1}+\cdots+d\lambda_{i_p}\big)^2 + 2\big(d\lambda_{i_1}+\cdots+d\lambda_{i_p}\big)d\ln(Z)+d\ln(Z)^2\big]$$

$$= \frac{1}{4}[\langle\widetilde{\psi}^\dagger dD^2\widetilde{\psi}\rangle + 2\langle\widetilde{\psi}^\dagger dD\widetilde{\psi}\rangle d\ln(Z)+d\ln(Z)^2]$$

$$= \frac{1}{4}\{\mathrm{Tr}[g(SdDS^{-1})^2] + 2\mathrm{Tr}(gSdDS^{-1})d\ln(Z)+d\ln(Z)^2\}.$$

But,

$$\frac{1}{2}d\ln(Z) = \mathrm{Tr}(gK) = n_i B_{ii} = -\frac{1}{2}\sum_{i=1}^{N} n_i d\lambda_i = -\frac{1}{2}\mathrm{Tr}(gSdDS^{-1}),$$

therefore,

$$\frac{1}{4}\sum_I \frac{dp_I^2}{p_I} = \frac{1}{4}\{\mathrm{Tr}[g(SdDS^{-1})^2] - [\mathrm{Tr}(gSdDS^{-1})]^2\} = \frac{1}{4}\sum_{i=1}^{N} n_i(1-n_i)d\lambda_i^2.$$

Also notice that this corresponds to evaluating the metric when $S^{-1}dS \equiv 0$. The nonclassical part is therefore given by

$$\sum_{i\neq j} n_i(1-n_j)|(S^{-1}dS)_{ij}|^2 \tanh^2\left(\frac{\lambda_i-\lambda_j}{2}\right).$$

1. Two-level case

We now suppose $N=2$ and without loss of generality we can write

$$H = x^\mu\sigma_\mu = (U\sigma_3U^{-1})r,\ \text{with } r=\sqrt{x^\mu x^\mu}.$$

The eigenstates of H are coherent states of SU(2), namely if we take $n^\mu = x^\mu/r$ and $z=(n^1+in^2)/(1+n^3)$ is the stereographic projection with respect to the south pole of the sphere, we can write

$$|z\rangle = \frac{1}{(1+|z|^2)^{1/2}}\begin{pmatrix}1\\ z\end{pmatrix}$$

satisfies $H|z\rangle = r|z\rangle$. We can then write

$$U = \frac{1}{(1+|z|^2)^{1/2}}\begin{pmatrix}1 & -\bar{z}\\ z & 1\end{pmatrix},$$

clearly,

$$(U^{-1}dU)_{21} = \frac{dz}{(1+|z|^2)},$$

and

$$|(U^{-1}dU)_{21}|^2 = \frac{dzd\bar{z}}{(1+|z|^2)^2} = \frac{1}{4}d\Omega^2,$$

where $d\Omega^2 = \delta_{\mu\nu}dn^\mu dn^\nu$ is the round metric on the sphere. Denoting $n \equiv n_1 = 1/(e^{-r}+1)$ we have $n_2 = 1-n_1$, so

$$ds^2 = \tfrac{1}{4}\tanh^2(r)[n^2+(1-n)^2]d\Omega^2 + \tfrac{1}{2}n(1-n)dr^2.$$

Also

$$n(1-n) = \frac{1}{(e^{-r}+1)(e^r+1)} = \frac{1}{2[\cosh(r)+1]},$$

and

$$n^2+(1-n)^2 = \frac{(e^r+1)^2+(e^{-r}+1)^2}{(e^r+1)^2(e^{-r}+1)^2} = \frac{1}{2}\frac{\cosh(r)}{\cosh^2\left(\frac{r}{2}\right)} = \frac{\cosh(r)}{1+\cosh(r)}$$

so

$$\begin{aligned}\frac{1}{4}\tanh^2(r)[n^2+(1-n)^2]d\Omega^2 &= \frac{1}{4}\frac{\cosh(r)}{1+\cosh(r)}\tanh^2(r)d\Omega^2\\ &= \frac{1}{4}\frac{\sinh^2(r)}{[1+\cosh(r)]\cosh(r)}d\Omega^2\\ &= \frac{1}{4}\frac{\cosh(r)-1}{\cosh(r)}d\Omega^2.\end{aligned}$$

Therefore,

$$ds^2 = \frac{1}{4}\left(\frac{\cosh(r)-1}{\cosh(r)}d\Omega^2 + \frac{dr^2}{\cosh(r)+1}\right).$$

2. Thermal states

With the replacement $x^\mu\sigma_\mu \to \beta E n^\mu\sigma_\mu$, we get

$$\begin{aligned}ds^2 &= \frac{1}{4}\left(\frac{\cosh(\beta E)-1}{\cosh(\beta E)}d\Omega^2 + \frac{d(\beta E)^2}{\cosh(\beta E)+1}\right)\\ &= \frac{1}{4}\left(\frac{\cosh(\beta E)-1}{\cosh(\beta E)}d\Omega^2 + \frac{\beta^2dE^2}{\cosh(\beta E)+1} + 2\frac{\beta E d\beta dE}{\cosh(\beta E)+1} + \frac{E^2d\beta^2}{\cosh(\beta E)+1}\right).\end{aligned}$$

Notice that since $ds^2 = \langle(\psi^\dagger K\psi)^2\rangle - \langle\psi^\dagger K\psi\rangle^2$, we have that a variation along β produces

$$\psi^\dagger K\psi = -\tfrac{1}{2}d\beta\psi^\dagger H\psi,$$

and hence

$$\begin{aligned}\langle(\psi^\dagger K\psi)^2\rangle - \langle\psi^\dagger K\psi\rangle^2 &= \frac{1}{4}d\beta^2[\langle(\psi^\dagger H\psi)^2\rangle - \langle\psi^\dagger H\psi\rangle] = \frac{1}{4}d\beta^2\frac{\partial^2\ln Z}{\partial\beta^2}\\ &= -\frac{1}{4}d\beta^2\frac{\partial\langle\psi^\dagger H\psi\rangle}{\partial\beta} = -d\beta^2\frac{1}{4}\frac{\partial T}{\partial\beta}\frac{\partial E}{\partial T}\\ &= d\beta^2\frac{1}{4}T^2\frac{\partial E}{\partial T} = d\beta^2\frac{T^2}{4}c,\end{aligned}$$

where $c = \partial E/\partial T$ is the specific heat of the system.

Now if we have a Chern insulator in two dimensions as in the main text, we will have $\mathcal{H} = \int_{\mathrm{BZ}} d^2k/(2\pi)^2\psi^\dagger_{\mathbf{k}}H(\mathbf{k})\psi_{\mathbf{k}}$ with $H(\mathbf{k}) = d^\mu(\mathbf{k})\sigma_\mu = E(\mathbf{k})n^\mu(\mathbf{k})\sigma_\mu$ depending on some parameters $\{t_i\}$, and the Bures metric will be

$$\begin{aligned}ds^2 = \int_{\mathrm{BZ}}\frac{d^2k}{(2\pi)^2}\Bigg[\frac{1}{4}\Bigg(&\frac{\cosh[\beta E(\mathbf{k})]-1}{\cosh[\beta E(\mathbf{k})]}\delta_{\mu\nu}\frac{\partial n^\mu(\mathbf{k})}{\partial t_i}\frac{\partial n^\mu(\mathbf{k})}{\partial t_j}dt_idt_j\\ &+ \frac{E^2(\mathbf{k})d\beta^2}{\cosh[\beta E(\mathbf{k})]+1} + 2\frac{\beta E(\mathbf{k})\frac{\partial E(\mathbf{k})}{\partial t_i}dt_id\beta}{\cosh[\beta E(\mathbf{k})]+1} + \frac{\beta^2\frac{\partial E(\mathbf{k})}{\partial t_i}\frac{\partial E(\mathbf{k})}{\partial t_j}dt_idt_j}{\cosh[\beta E(\mathbf{k})]+1}\Bigg)\Bigg].\end{aligned}$$

We remark that for the case of the topological superconductor considered in the main text the charge is only conserved modulo 2. However, since the expression for the fidelity is, apart from an overall square root, the same, as a function of $\vec{d}$ and β, as the

charge conserving ones (see Appendix A), the resulting susceptibility has the same functional form, apart from an overall factor of 1/2 (this follows from $\sqrt{1-x} \approx 1 - x/2$ for small x). Related to this derivation, we refer the reader to Ref. [27] to a recent derivation of the symmetric logarithmic derivative of fermionic Gaussian states by Carollo *et al.*

[1] L. D. Landau, On the theory of phase transitions, Zh. Eksp. Teor. Fiz. **7**, 19 (1937).

[2] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Quantized Hall Conductance in a Two-Dimensional Periodic Potential, Phys. Rev. Lett. **49**, 405 (1982).

[3] M. V. Berry, Quantal phase factors accompanying adiabatic changes, Proc. R. Soc. London, Ser. A **392**, 45 (1984).

[4] Y. Ando, Topological insulator materials, J. Phys. Soc. Jpn. **82**, 102001 (2013).

[5] B. Mera, Topological response of gapped fermions to a U(1) gauge field, arXiv:1705.04394.

[6] S. Ryu and Y. Hatsugai, Topological Origin of Zero-Energy Edge States in Particle-Hole Symmetric Systems, Phys. Rev. Lett. **89**, 077002 (2002).

[7] S. Sachdev, *Quantum Phase Transitions* (Wiley, Hoboken, NJ, 2007).

[8] G. F. A. Osterloh, L. Amico, and R. Fazio, Scaling of entanglement close to a quantum phase transition, Nature (London) **416**, 608 (2002).

[9] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Entanglement in Quantum Critical Phenomena, Phys. Rev. Lett. **90**, 227902 (2003).

[10] S.-Q. Su, J.-L. Song, and S.-J. Gu, Local entanglement and quantum phase transition in a one-dimensional transverse field Ising model, Phys. Rev. A **74**, 032308 (2006).

[11] T. P. Oliveira and P. D. Sacramento, Entanglement modes and topological phase transitions in superconductors, Phys. Rev. B **89**, 094512 (2014).

[12] A. Uhlmann, The transition probability in the state space of a C*-algebras, Rep. Math. Phys. **9**, 273 (1976).

[13] P. M. Alberti, A note on the transition probability over C*-algebras, Lett. Math. Phys. **7**, 25 (1983).

[14] H. T. Quan, Z. Song, X. F. Liu, P. Zanardi, and C. P. Sun, Decay of Loschmidt Echo Enhanced by Quantum Criticality, Phys. Rev. Lett. **96**, 140604 (2006).

[15] P. Zanardi and N. Paunković, Ground state overlap and quantum phase transitions, Phys. Rev. E **74**, 031123 (2006).

[16] N. Paunković, P. D. Sacramento, P. Nogueira, V. R. Vieira, and V. K. Dugaev, Fidelity between partial states as a signature of quantum phase transitions, Phys. Rev. A **77**, 052302 (2008).

[17] P. D. Sacramento, N. Paunković, and V. R. Vieira, Fidelity spectrum and phase transitions of quantum systems, Phys. Rev. A **84**, 062318 (2011).

[18] B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Uhlmann Connection in Fermionic Systems Undergoing Phase Transitions, Phys. Rev. Lett. **119**, 015702 (2017).

[19] B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Boltzmann-Gibbs states in topological quantum walks and associated many-body systems: fidelity and Uhlmann parallel transport analysis of phase transitions, J. Phys. A **50**, 365302 (2017).

[20] Let us stress that the Cauchy-Schwarz inequality, in general, reads $|\langle A, B\rangle| \leqslant ||A||||B||$, and here we have considered $A = |\sqrt{\rho_1}\sqrt{\rho_2}|^{1/2}$ and $B = |\sqrt{\rho_1}\sqrt{\rho_2}|^{1/2} V U_1 U_2^\dagger$ to obtain the upper bound of $\mathrm{Re}\langle w_1, w_2\rangle$. Notice that if we had considered $A = |\sqrt{\rho_1}\sqrt{\rho_2}|$ and $B = V U_1 U_2^\dagger$, then we would have had an extra multiplicative factor, the trace of the identity (i.e., the dimension of the Hilbert space).

[21] N. Paunković and V. Rocha Vieira, Macroscopic distinguishability between quantum states defining different phases of matter: Fidelity and the Uhlmann geometric phase, Phys. Rev. E **77**, 011129 (2008).

[22] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, 1st ed., Cambridge Series on Information and the Natural Sciences (Cambridge University Press, Cambridge, UK, 2000).

[23] L. Campos Venuti and P. Zanardi, Quantum Critical Scaling of the Geometric Tensors, Phys. Rev. Lett. **99**, 095701 (2007).

[24] P. Zanardi, P. Giorda, and M. Cozzini, Information-Theoretic Differential Geometry of Quantum Phase Transitions, Phys. Rev. Lett. **99**, 100603 (2007).

[25] G. Palumbo and N. Goldman, Revealing Tensor Monopoles Through Quantum-Metric Measurements, Phys. Rev. Lett. **121**, 170401 (2018).

[26] P. Zanardi, M. G. A. Paris, and L. Campos Venuti, Quantum criticality as a resource for quantum estimation, Phys. Rev. A **78**, 042105 (2008).

[27] A. Carollo, B. Spagnolo, and D. Valenti, Symmetric logarithmic derivative of fermionic gaussian states, Entropy **20**, 485 (2018).

[28] O. Viyuela, A. Rivas, S. Gasparinetti, A. Wallraff, S. Filipp, and M. A. Martin-Delgado, Observation of topological Uhlmann phases with superconducting qubits, npj Quantum Information **4**, 10 (2018).

[29] L. Leonforte, D. Valenti, B. Spagnolo, and A. Carollo, Uhlmann number in translational invariant systems, arXiv:1806.08592.

[30] B. A. Bernevig and T. L. Hughes, *Topological Insulators and Topological Superconductors* (Princeton University Press, Princeton, NJ, 2013).

[31] G. E. Volovik, Fermion zero modes on vortices in chiral superconductors, J. Exp. Theor. Phys. Lett. **70**, 609 (1999).

[32] N. Read and D. Green, Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum Hall effect, Phys. Rev. B **61**, 10267 (2000).

[33] D. A. Ivanov, Non-Abelian Statistics of Half-Quantum Vortices in p-Wave Superconductors, Phys. Rev. Lett. **86**, 268 (2001).

[34] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Two-Dimensional Density-Matrix Topological Fermionic Phases: Topological Uhlmann Numbers, Phys. Rev. Lett. **113**, 076408 (2014).

[35] B. Mera, C. Vlachou, N. Paunković, V. R. Vieira, and O. Viyuela, Dynamical phase transitions at finite temperature from fidelity and interferometric Loschmidt echo induced metrics, Phys. Rev. B **97**, 094110 (2018).

[36] D. Sticlet, F. Piéchon, J. N. Fuchs, P. Kalugin, and P. Simon, Geometrical engineering of a two-band Chern insulator in two dimensions with arbitrary topological index, Phys. Rev. B **85**, 165456 (2012).

[37] A. Rivas, O. Viyuela, and M. A. Martin-Delgado, Density-matrix Chern insulators: Finite-temperature generalization of topological insulators, Phys. Rev. B **88**, 155141 (2013).

[38] A. Kitaev, Fault-tolerant quantum computation by anyons, Ann. Phys. **303**, 2 (2003).

[39] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, Non-Abelian anyons and topological quantum computation, Rev. Mod. Phys. **80**, 1083 (2008).

[40] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, Quantum memories at finite temperature, Rev. Mod. Phys. **88**, 045005 (2016).

Classical and Quantum
Gravity

PURPOSE-LED
PUBLISHING™

PAPER

Gauge protected entanglement between gravity and matter

To cite this article: N Paunković and M Vojinović 2018 *Class. Quantum Grav.* **35** 185015

View the article online for updates and enhancements.

You may also like

- CEWQO Topical Issue
Mirjana Bozic and Margarita Man'ko
- How to control decoherence and entanglement in quantum complex systems?
V M Akulin, G Kurizki and D A Lidar
- Effect of depolarizing noise on entangled photons
Ruchipas Bavontaweepanya

This content was downloaded from IP address 147.91.81.84 on 19/09/2025 at 14:32

IOP Publishing Classical and Quantum Gravity
Class. Quantum Grav. **35** (2018) 185015 (27pp) https://doi.org/10.1088/1361-6382/aad7f1

Gauge protected entanglement between gravity and matter

N Paunković[1,2,3] **and M Vojinović**[4]

[1] Instituto de Telecomunicaẽoes, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal
[2] Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal
[3] CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal
[4] Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

E-mail: npaunkov@math.tecnico.ulisboa.pt and vmarko@ipb.ac.rs

Received 14 May 2018, revised 1 August 2018
Accepted for publication 3 August 2018
Published 29 August 2018

Abstract

We show that, as a consequence of the local Poincaré symmetry, gravity and matter fields have to be entangled, unless the overall action is carefully fine-tuned. First, we present a general argument, applicable to any particular theory of quantum gravity with matter, by performing the analysis in the abstract nonperturbative canonical framework, demonstrating the nonseparability of the scalar constraint, thus promoting the entangled states as the physical ones. Also, within the covariant framework, using a particular toy model, we show explicitly that the Hartle–Hawking state in the Regge model of quantum gravity is entangled. Our result is potentially relevant for the quantum-to-classical transition, taken within the framework of the decoherence programme: due to the gauge symmetry requirements, the matter does not decohere, it is by default decohered by gravity. Generically, entanglement is a consequence of interaction. This new entanglement could potentially, in form of an 'effective interaction', bring about corrections to the weak equivalence principle, further confirming that spacetime as a smooth four-dimensional manifold is an emergent phenomenon. Finally, the existence of the gauge-protected entanglement between gravity and matter could be seen as a criterion for a plausible theory of quantum gravity, and in the case of perturbative quantisation approaches, a confirmation of the persistence of the manifestly broken gauge symmetry.

Keywords: quantum entanglement, quantum gravity, diffeomorphism invariance, scalar constraint

 © 2018 IOP Publishing Ltd Printed in the UK

1. Introduction

The unsolved problems of formulating quantum theory of gravity (QG) and interpreting quantum mechanics (QM) are arguably the two most prominent ones of the modern theoretical physics. So far, most of the approaches to solve the two were studied independently. Indeed, the majority of the interpretations of QM do not involve explicit dynamical effects (with notable exceptions of the spontaneous collapse and the de Broglie–Bohm theories), while the researchers from the QG community often just adopt some particular interpretation of QM, assuming that it contains no unresolved issues. Nevertheless, the two problems share a number of similar unsolved questions and counter-intuitive features. A prominent example is nonlocality: entanglement-based nonlocality in QM, as well as the anticipated explicit dynamical nonlocality in QG (a consequence of quantum superpositions of different gravitational fields, i.e. different spacetimes and their respective causal orders). Another prominent issue relevant for both standard QM and QG is the quantum-to-classical transition and the related measurement problem.

In relation to the latter, decoherence is in QM the standard approach to the emergence of classicality: due to huge complexity of macroscopic ('classical') systems and the surrounding environment (bath), the (*for all practical purposes*) inevitable interaction between the two leads to the entanglement and the loss of coherence. While technically this is completely within the standard QM, when coupled with additional assumptions, such as the many-world interpretation (likely to be the predominant within the community working on decoherence and quantum-to-classical transition), the decoherence offers a possible solution to the measurement problem. In an alternative approach, problems with quantising gravity led to the half century old idea of gravitationally induced objective collapse of the wave function [1] (for an overview, see for example [2], chapter III.B): roughly speaking, due to the position uncertainty of massive bodies, which are the sources of gravitational field, the latter exhibits quantum fluctuations that decohere the matter, forcing it (or, rather both the matter and gravity) to collapse in a well defined (classical) state. Without invoking objective collapse, decoherence of quantum matter by purely classical gravity was studied in [3, 4]. In the context of perturbative quantum gravity, the topic of gravitationally induced decoherence of matter, taken purely within the scope of standard QM (i.e. in the same fashion in which macroscopic bodies decohere due to inevitable interaction with surrounding photons, neutrinos, microwave background radiation, etc), became recently an intensive field of research [5], see also [6] and the references therein for decoherence in the context of cosmological inflation. In addition, a lot of research focuses on entanglement induced by the presence of horizons in curved spacetime, in approaches based on the holography conjecture and in the studies of the black hole information problem [7] (for a review, see recent lecture notes [8]). In particular, these approaches study the entanglement between the degrees of freedom (*both* gravitational and matter) on the two sides of the horizon.

In this paper we study the entanglement between gravitational and matter fields, in the context of an abstract nonperturbative theory of quantum gravity, as well as on the example of the Hartle–Hawking state in the Regge quantum gravity model, and show that the two fields should *always* be entangled. Our approach is different from the standard one, studied in the perturbative framework: instead of 'for all practical purposes' inevitable fast interaction-induced decoherence from initially product states between two sub-systems [5, 9–13], we show that the gauge symmetry requirements (coming in particular from the local Poincaré symmetry) secure the entangled states between matter and gravity as physical states. We call the latter the *gauge-protected* decoherence, in contrast to the *dynamical* decoherence of the former. In addition, unlike the horizon-based studies, we discuss the entanglement between

the gravitational and the matter degrees of freedom, rather than between the two specially chosen regions of spacetime.

Our analysis rests on two main assumptions. First, we assume the validity of the local Poincaré symmetry at the quantum level. In the classical field theory, the local Poincaré symmetry is a formalisation of the principle of general relativity, which is one of the foundational principles of Einstein's theory of gravity. It is therefore natural to assume that this gauge symmetry exists at the quantum level as well. Second, at the classical level we assume the validity of the equivalence principle, which is also the main ingredient of Einstein's general relativity. In particular, we assume its 'strong' version, namely that the equivalence principle applies to all matter fields (i.e. all non-gravitational fields) present in nature.

Given these two assumptions, we focus on the general nonperturbative abstract canonical quantisation of the gravitational and matter fields, thus giving a generic model-independent argument for a theory of quantum gravity with matter. We analyse the consequences of the local Poincaré symmetry-enforced scalar, 3-diffeomorphism and local Lorentz constraints on the structure of the total Hilbert space of the theory. Namely, since the physical states must be invariant with respect to the gauge symmetry, the constraints induce the Gupta–Bleuler-like conditions on the state vectors. Based on the equivalence principle, we then show that the particular non-separable form of the scalar constraint renders typical product states non-invariant. Thus, it eliminates the product states from the physical Hilbert space of the quantum theory, unless the interaction between gravity and matter is specifically designed to circumvent the non-invariance of product states. In this way, the local Poincaré symmetry protects the existence of entanglement between the gravitational and matter fields.

In order to verify our results obtained within the abstract canonical framework, we also study the covariant (i.e. path integral) quantisation. In particular, knowing that the Hartle–Hawking state [14] satisfies the scalar constraint, and is therefore an element of the physical Hilbert space, we explicitly test whether the matter and gravitational fields are entangled for this state vector. We perform the calculation in the Regge quantum gravity model, since it is one of the simplest models which provide an explicit definition of the gravitational path integral with matter, and show that the gravitational and matter fields are indeed entangled for the Hartle–Hawking state constructed on a simple toy example triangulation.

Therefore, our analysis shows that either gravity and matter fields are indeed entangled, or there exists an additional, unknown property of the action, implementing the fine tuning needed to allow for the invariance of separable states.

The paper is organised as follows. Section 2 is divided into three subsections. The first is devoted to the recapitulation of the Hamiltonian structure of Poincaré gauge theories. The second outlines the procedure of nonperturbative canonical quantisation of constrained systems and its application to the case of gravity with matter fields. In the third subsection we use those results to show that the scalar constraint suppresses the existence of separable states of a matter-gravity system. In section 3, we present a standard entanglement criterion for pure bipartite quantum states and discuss it, within the framework of the path integral quantisation, for the case of the Hartle–Hawking state of quantum fields of gravity and matter. In section 4, we first introduce the Regge model of quantum gravity, and then apply it to evaluate the entanglement criterion for the Hartle–Hawking state, demonstrating that gravity and matter are indeed entangled in this state. Finally, in section 5 we present the summary of the results, their discussion, and possible future lines of research.

It is important to stress that the gauge-protected entanglement is not an automatic consequence of the universal coupling between gravity and matter, or the fact that matter fields are always defined over some background spacetime geometry. For example, in perturbative gravity approach, it is quite possible to write the separable state between gravity and matter as

$$|\Psi\rangle = |g\rangle \otimes |\phi\rangle,$$

where $|g\rangle$ is the graviton state vector, while $|\phi\rangle$ is the state vector of a scalar particle (both vectors obtained by acting with graviton and scalar creation operators on the Minkowski vacuum state $|0\rangle \equiv |0\rangle_G \otimes |0\rangle_M$). The reason why such a state can be considered legitimate is that local Poincaré symmetry is explicitly broken in the perturbative gravity approach, with both matter and gravity being treated as spin-zero and spin-two fields, respectively, living on a Minkowski spacetime manifold. A similar situation arises in perturbative string theory, where local Poincaré symmetry is also manifestly broken. However, in quantum gravity models where the local Poincaré symmetry is not violated, our analysis shows that a generic product state between gravity and matter would fail to be gauge invariant. Thus, the gauge-protected entanglement between gravity and matter is a nontrivial statement and a consequence of local Poincaré symmetry, rather than an automatic property of matter fields living on a spacetime manifold.

Our notation and conventions are as follows. We will work in the natural system of units in which $c = \hbar = 1$ and $G = l_p^2$, where l_p is the Planck length. By convention, the metric of spacetime will have the spacelike Lorentz signature $(-,+,+,+)$. The spacetime indices are denoted with lowercase Greek letters $\mu, \nu, \ldots$ and take the values $0, 1, 2, 3$. The spatial part of these, taking values $1, 2, 3$, will be denoted with lowercase Latin letters $i, j, \ldots$ from the middle of the alphabet. The $SO(3,1)$ group indices will be denoted with the lowercase Latin letters $a, b, \ldots$ from the beginning of the alphabet, and take the values $0, 1, 2, 3$. The Lorentz-invariant metric tensor is denoted as η_{ab}. The capital Latin indices $A, B, \ldots$ count the field components in a particular representation of the $SO(3,1)$ group, and take the values from 1 up to the dimension of that representation. Quantum operators will always carry a hat, $\hat{\phi}(x)$, $\hat{g}(x)$, etc. Finally, we will systematically denote the values of functions with parentheses, $f(x)$, while functionals will be denoted with brackets, $F[\phi]$.

2. Entanglement from the scalar constraint

This section is dedicated to the analysis of the constraints imposed by the relativity and equivalence principles. In section 2.1 we briefly recapitulate the classical Hamiltonian structure of gravitational interaction, followed by a short review of canonical quantisation, presented in section 2.2. After that, in section 2.3 we present the main result of our paper: we show that the scalar constraint, and possibly the 3-diffeomorphism constraint, bring about the generic entanglement between gravity and matter.

2.1. Hamiltonian structure of Poincaré gauge theories

We begin with a short review of the Hamiltonian structure of gravitational interaction, based on the local Poincaré symmetry. This subsection is aimed to be only a review of the main results, so we will skip all proofs and derivations. The details of the Hamiltonian structure for Poincaré gauge theories (PGT) can be found in many textbooks, see for example [15], chapter V, and the references therein.

We will assume a foliation of spacetime into space and time, with the spacetime topology $\mathcal{M}_4 = \Sigma_3 \times \mathbb{R}$, where Σ_3 is the $3D$ hypersurface. For the purpose of generality, we will describe the gravitational field as $g(x)$ and matter fields as $\phi(x)$, without specifying their exact field content, except in examples. A typical example would be the Einstein–Cartan gravity coupled to a Dirac matter field, so that the choice of fundamental gravitational fields g would

be the tetrads $e^a{}_\mu(x)$ and the spin connection $\omega^{ab}{}_\mu(x)$, while the choice for the fundamental matter field ϕ would be a Dirac fermion field $\psi(x)$. However, other choices for g and ϕ are also possible, for example the metric tensor $g_{\mu\nu}$ for gravity and the electromagnetic potential A^μ for matter, etc. Since our analysis is largely independent of such choices, we will stick to the abstract notation g and ϕ, assuming that one can apply our analysis to each particular concrete choice of fundamental fields.

Given the above notation, we will assume that the action of the theory can be written as

$$S[g,\phi] = S_G[g] + S_M[g,\phi], \tag{1}$$

where $S_G[g]$ is the action of the pure gravitational field, while $S_M[g,\phi]$ is the action of the matter fields coupled to gravity. Since the spacetime metric must both be a function of the gravitational field g and is always present in the definition of the dynamics of matter fields, the action for the matter fields cannot contain terms independent of g. This elementary fact is the crux of our main argument below, and is justified by the equivalence principle, which dictates how matter couples to gravity.

To a large extent, we also do not need to specify the details of the actions $S_G[g]$ and $S_M[g,\phi]$. We will only assume that the action (1) belongs to the PGT class of theories, i.e. that it is invariant with respect to local Poincaré group $P(4) = \mathbb{R}^4 \ltimes SO(3,1)$. Every theory belonging to the PGT class has the Hamiltonian with the following general structure [15]:

$$H = \int_{\Sigma_3} \mathrm{d}^3\vec{x}\left[N\mathcal{C} + N^i\mathcal{C}_i + N^{ab}\mathcal{C}_{ab}\right], \tag{2}$$

up to a 3-divergence. Here N, N^i and N^{ab} are Lagrange multipliers, the first two of which are commonly known as lapse and shift functions. The quantities $\mathcal{C}$, $\mathcal{C}_i$ and $\mathcal{C}_{ab}$ are usually known as the scalar constraint, 3-diffeomorphism constraint, and the local Lorentz constraint (sometimes also called the Gauss constraint), respectively. They are a (g,ϕ)-field representation of the 10 generators of the Poincaré group $P(4)$, in particular the time translation generator, the three space translation generators, and six local Lorentz generators (rotations and boosts). Note that the Hamiltonian (2) is always a linear combination of these constraints.

The constraints in (2) have the structure similar to the structure of the gravity-matter action (1), namely

$$\begin{aligned}
\mathcal{C} &= \mathcal{C}^G(g,\pi_g) + \mathcal{C}^M(g,\pi_g,\phi,\pi_\phi),\\
\mathcal{C}_i &= \mathcal{C}^G_i(g,\pi_g) + \mathcal{C}^M_i(g,\pi_g,\phi,\pi_\phi),\\
\mathcal{C}_{ab} &= \mathcal{C}^G_{ab}(g,\pi_g) + \mathcal{C}^M_{ab}(g,\pi_g,\phi,\pi_\phi),
\end{aligned} \tag{3}$$

where π_g and π_ϕ are the momenta canonically conjugated to the fields g and ϕ, respectively, defined as functional derivatives of the action with respect to the time-derivatives of the fields,

$$\pi_g(x) = \frac{\delta S}{\delta\partial_0 g(x)}, \qquad \pi_\phi(x) = \frac{\delta S}{\delta\partial_0\phi(x)}.$$

The general dependence (3) on the fields and momenta reflects the corresponding dependence in (1).

The exact forms of the gravitational terms of the constraints, namely $\mathcal{C}^G$, $\mathcal{C}^G_i$ and $\mathcal{C}^G_{ab}$, will be immaterial for our main argument presented in the section 2.3 below. In contrast, the structure of the matter terms $\mathcal{C}^M$, $\mathcal{C}^M_i$ and $\mathcal{C}^M_{ab}$ will be crucial, so we discuss it here in more detail. Choose a matter field such that it transforms according to some specific irreducible transformation of the Poincaré group, and denote it as $\phi^A(x)$, where the capital index A counts the field

components in that representation. Then the 3-diffeo constraint $\mathcal{C}_i^M$ and the Gauss constraint $\mathcal{C}_{ab}^M$ are given as

$$\begin{aligned} \mathcal{C}_i^M(g,\pi_g,\phi,\pi_\phi) &= \pi_{\phi A}\nabla_i{}^A{}_B\phi^B, \\ \mathcal{C}_{ab}^M(g,\pi_g,\phi,\pi_\phi) &= \pi_{\phi A}(M_{ab})^A{}_B\phi^B, \end{aligned} \tag{4}$$

where $\nabla_i{}^A{}_B$ is a covariant derivative for the irreducible representation according to which the field ϕ transforms, while $(M_{ab})^A{}_B$ is the representation of the generator M_{ab} of the Lorentz group $SO(3,1)$ in the same representation. In general, the covariant derivative depends on the spacetime metric or connection, which is a function of the fundamental gravitational fields g, and possibly their momenta π_g. The Lorentz group generators, on the other hand, do not depend on the spacetime geometry, so the Gauss constraint is actually independent of g and π_g, and we can write $\mathcal{C}_{ab}^M(g,\pi_g,\phi,\pi_\phi) = \mathcal{C}_{ab}^M(\phi,\pi_\phi)$.

In order to illustrate the two constraints, we will write (4) for the scalar and Dirac fields, as the most elementary examples. In the case of the scalar field, we write $\phi^A(x) = \varphi(x)$, where the index A takes only a single value. The covariant derivative acts on the scalar field as an ordinary derivative, while the representation of the Lorentz generators is trivial, so we can write

$$\mathcal{C}_i^M(\varphi,\pi_\varphi) = \pi_\varphi\partial_i\varphi, \qquad \mathcal{C}_{ab}^M(\varphi,\pi_\varphi) = \pi_\varphi\varphi. \tag{5}$$

We see that in the case of the scalar field, both constraints are independent of the gravitational fields and their momenta. In the case of the Dirac fields, we write $\phi^A(x) = (\psi^A(x), \bar\psi^A(x))$, where the index A now represents the spinorial index, and we will omit writing it. The covariant derivative acts on the Dirac field in the standard way,

$$\begin{aligned} \overrightarrow{\nabla}_\mu\psi &\equiv \partial_\mu\psi + \frac{1}{2}\omega^{ab}{}_\mu\sigma_{ab}\psi, \\ \bar\psi\overleftarrow{\nabla}_\mu &\equiv \partial_\mu\bar\psi - \frac{1}{2}\omega^{ab}{}_\mu\bar\psi\sigma_{ab}, \end{aligned} \tag{6}$$

where $\omega^{ab}{}_\mu$ is the spin connection, $\sigma_{ab} = \frac{1}{4}[\gamma_a,\gamma_b]$, and γ_a are the standard Dirac gamma-matrices satisfying the anticommutation relation $\{\gamma_a,\gamma_b\} = -2\eta_{ab}$. The representation of the Lorentz generators for the case of the Dirac field is $M_{ab} = \sigma_{ab}$. Denoting the conjugate momentum for ψ as $\bar\pi$ and conjugate momentum for $\bar\psi$ as π, we can write the constraints (4) as:

$$\begin{aligned} \mathcal{C}_i^M(\omega,\psi,\bar\pi,\bar\psi,\pi) &= \bar\pi\overrightarrow{\nabla}_i\psi + (\bar\psi\overleftarrow{\nabla}_i)\pi, \\ \mathcal{C}_{ab}^M(\psi,\bar\pi,\bar\psi,\pi) &= \bar\pi\sigma_{ab}\psi - \bar\psi\sigma_{ab}\pi. \end{aligned} \tag{7}$$

Note that here, unlike in the scalar field example, the 3-diffeo constraint contains the spin connection $\omega^{ab}{}_\mu$, which is a part of the gravitational field $g = (e^a{}_\mu, \omega^{ab}{}_\mu)$ for the Einstein–Cartan gravity.

In contrast to the 3-diffeo and Gauss constraints (4), the scalar constraint $\mathcal{C}^M$ has a more complicated form,

$$\mathcal{C}^M(g,\pi_g,\phi,\pi_\phi) = \pi_{\phi A}\nabla_\perp{}^A{}_B\phi^B - \frac{1}{N}\mathcal{L}_M(g,\pi_g,\phi,\pi_\phi), \tag{8}$$

where the matter Lagrangian density is defined via

$$S_M[g,\phi] = \int \mathrm{d}^4x\,\mathcal{L}_M(g,\partial g,\phi,\partial\phi),$$

and $\nabla_\perp \equiv n^\mu \nabla_\mu$ is the covariant derivative in the direction of the timelike vector n^μ orthogonal to the spacelike hypersurface Σ_3. The vector n^μ obviously depends on the spacetime metric $g_{\mu\nu}$, and is thus a function of the fundamental gravitational fields g.

There are several things to note regarding the scalar constraint (8). First, it is clear that $N\mathcal{C}^M$ is the Legendre transformation of the Lagrangian density $\mathcal{L}_M$ with respect to the 'velocity' $N\nabla_\perp\phi$. Second, in contrast to the constraints (4), which depend only on the symmetry transformation properties of the fields, the form of the scalar constraint (8) depends also on the choice of the matter Lagrangian density $\mathcal{L}_M$, and is therefore described by the dynamics of the matter fields coupled to gravity. And third, the scalar constraint $\mathcal{C}^M$ always necessarily depends on the gravitational fields g, in contrast to the 3-diffeo constraint which may or may not depend on g, and the Gauss constraint which never depends on g. As we already suggested above, this is because the Lagrangian of the matter fields coupled to gravity always contains the gravitational degrees of freedom, courtesy of the equivalence principle.

Let us illustrate this dependence of $\mathcal{C}^M$ on the gravitational fields g in the case of the Dirac field. The action for the Dirac field $\phi = (\psi, \bar\psi)$ coupled to the gravitational fields $g = (e^a{}_\mu, \omega^{ab}{}_\mu)$ is given as

$$S_M[e,\omega,\psi,\bar\psi] = \int \mathrm{d}^4x\, e \left(\frac{\mathrm{i}}{2}\bar\psi\gamma^a e^\mu{}_a \overset{\leftrightarrow}{\nabla}_\mu \psi - m\bar\psi\psi \right), \tag{9}$$

where e is the determinant of the tetrad $e^a{}_\mu$, while $e^\mu{}_a$ is the inverse tetrad. In addition, $\overset{\leftrightarrow}{\nabla}_\mu \equiv \overset{\rightarrow}{\nabla}_\mu - \overset{\leftarrow}{\nabla}_\mu$, and the covariant derivatives $\overset{\rightarrow}{\nabla}_\mu$ and $\overset{\leftarrow}{\nabla}_\mu$ act to the right and to the left as defined in (6), from which one can see that the action also explicitly depends on the connection $\omega^{ab}{}_\mu$. From the action one can read off the Lagrangian density, and calculate the scalar constraint (8) as

$$\mathcal{C}^M(e,\omega,\psi,\bar\psi) = -\frac{e}{N}\left(\frac{\mathrm{i}}{2}\bar\psi\gamma^a e^\mu{}_a \left(\delta^\mu_\nu + n^\mu n_\nu\right) \overset{\leftrightarrow}{\nabla}_\nu \psi - m\bar\psi\psi \right).$$

Note that the quantity $\delta^\mu_\nu + n^\mu n_\nu$ is a projector to the hypersurface Σ_3.

2.2. Canonical quantisation

Having discussed the Hamiltonian structure of the action (1), we now pass on to a short description of the canonical quantisation of the theory. The quantisation of an arbitrary physical system with constraints is performed in the standard way, using the Dirac's procedure [16, 17] (see [15] for a review). One begins by classifying all constraints of the theory into the first and the second class. The second class constraints are then eliminated by passing from the Poisson brackets to the Dirac brackets. The first class constraints remain and represent the generators of the gauge symmetry. In general, the Hamiltonian of the theory can be written as

$$H = H_0 + \lambda^A \mathcal{C}_A, \tag{10}$$

where λ^A are Lagrange multipliers, $\mathcal{C}_A$ are first class constraints, and H_0 is the part of the Hamiltonian which describes the evolution of the physical degrees of freedom. Given all this, the quantisation is performed in the Heisenberg picture, promoting fundamental fields $\phi(x)$ to quantum mechanical operators $\hat\phi(x)$, and introducing the state vectors $|\Psi\rangle \in \mathcal{H}_{\text{kin}}$, where $\mathcal{H}_{\text{kin}}$ is the kinematical Hilbert space of the theory. The Dirac brackets between the fields and their momenta are then promoted to the commutators of the corresponding operators. The Hamiltonian, being a functional of the fields and momenta, also becomes an operator, providing the usual Heisenberg equations of motion for the field operators,

$$\mathrm{i}\frac{\partial\hat{\phi}(x)}{\partial t} = [\hat{\phi}(x), \hat{H}].$$

Finally, the kinematical Hilbert space $\mathcal{H}_{\text{kin}}$ is projected onto its gauge invariant subspace $\mathcal{H}_{\text{phys}}$, by requiring that every state vector $|\Psi\rangle \in \mathcal{H}_{\text{phys}}$ is annihilated by the generators of the gauge symmetry group,

$$\hat{\mathcal{C}}_A|\Psi\rangle = 0.$$

In quantum electrodynamics these conditions are known as Gupta–Bleuler quantisation conditions [18, 19]. This requirement ensures that the gauge symmetry of the classical theory remains to be a symmetry of the quantum theory as well.

Of course, one cannot hope to implement the above quantisation programme in full detail for the general action (1), especially without the detailed specification of the fundamental degrees of freedom that define the theory. Instead, we assume that the quantisation programme has been carried out in detail, and that all quantities we will write are well defined. This approach has one important feature and one important drawback. The feature is generality—our main argument for the inevitable entanglement between gravity and matter, to be presented in section 2.3, should hold for every particular quantum theory constructed in the above way, as it does not actually depend on the details of the quantisation. The drawback is abstractness—in using such a general formalism and making a flat assumption that all details are well defined, we lose the capability to provide any concrete examples. That said, in section 4 we discuss one rigorously defined example of a theory of quantum gravity with matter (Regge quantum gravity), and demonstrate the entanglement between gravity and matter fields. Unlike the canonical quantisation discussed in this section, that example will be done in the framework of the path integral quantisation.

Keeping this disclaimer in mind, we proceed along the lines outlined above and perform the canonical quantisation. The most prominent property of our model is the structure of the Hilbert space of the theory. The initial kinematical Hilbert space $\mathcal{H}_{\text{kin}} = \mathcal{H}_G \otimes \mathcal{H}_M$ has a natural product structure between the gravitational and matter Hilbert spaces, since we have two sets of fields, $\hat{g}$ and $\hat{\phi}$, corresponding to gravity and matter, respectively. Thus, we have a naturally preferred bipartite physical system, because gravitational and matter degrees of freedom can be fully distinguished from each other. Second, in order to preserve the Poincaré gauge symmetry of the theory at the quantum level, we have to pass from the kinematical Hilbert space to the gauge invariant, physical Hilbert space $\mathcal{H}_{\text{phys}}$. By definition, a state $|\Psi\rangle \in \mathcal{H}_{\text{kin}}$ is an element of $\mathcal{H}_{\text{phys}}$ iff it satisfies

$$\begin{aligned}\hat{\mathcal{C}}_{ab}|\Psi\rangle &\equiv \left[\mathcal{C}^G_{ab}(\hat{g},\hat{\pi}_g) + \mathcal{C}^M_{ab}(\hat{\phi},\hat{\pi}_\phi)\right]|\Psi\rangle = 0,\\ \hat{\mathcal{C}}_i|\Psi\rangle &\equiv \left[\mathcal{C}^G_i(\hat{g},\hat{\pi}_g) + \mathcal{C}^M_i(\hat{g},\hat{\pi}_g,\hat{\phi},\hat{\pi}_\phi)\right]|\Psi\rangle = 0,\end{aligned} \tag{11}$$

and

$$\hat{\mathcal{C}}|\Psi\rangle \equiv \left[\mathcal{C}^G(\hat{g},\hat{\pi}_g) + \mathcal{C}^M(\hat{g},\hat{\pi}_g,\hat{\phi},\hat{\pi}_\phi)\right]|\Psi\rangle = 0. \tag{12}$$

As stated above, we assume that the operators $\hat{\mathcal{C}}_{ab}$, $\hat{\mathcal{C}}_i$ and $\hat{\mathcal{C}}$ are well defined, that operator ordering choice has been fixed, as well as all other necessary technical choices, in order for the expressions above to make sense mathematically.

We argue that, due to these constraint equations, there are no states in $\mathcal{H}_{\text{phys}}$ which can be written as product states of the form $|\Psi_G\rangle \otimes |\Psi_M\rangle$, where $|\Psi_G\rangle \in \mathcal{H}_G$ and $|\Psi_M\rangle \in \mathcal{H}_M$, i.e. the states in $\mathcal{H}_{\text{phys}}$ are entangled. We focus on the scalar constraint (12), while the constraints

(11) are either irrelevant or redundant for our analysis. This main argument of our paper is presented in the next subsection.

2.3. Entanglement

Given a state vector $|\Psi\rangle \in \mathcal{H}_{\text{kin}} = \mathcal{H}_G \otimes \mathcal{H}_M$, it is an element of the physical Hilbert space $\mathcal{H}_{\text{phys}}$ if it satisfies the Gauss and 3-diffeo constraints (11) and the scalar constraint (12). Choosing the eigenbases $\{|g\rangle\}$ and $\{|\phi\rangle\}$ of the quantum field operators $\hat{g}$ and $\hat{\phi}$, respectively, we can work in the so-called field representation, defined as

$$\begin{aligned} \langle g|\hat{g} &= g\langle g|, & \langle g|\hat{\pi}_g &= -\mathrm{i}\frac{\delta}{\delta g}\langle g|, \\ \langle \phi|\hat{\phi} &= \phi\langle \phi|, & \langle \phi|\hat{\pi}_\phi &= -\mathrm{i}\frac{\delta}{\delta \phi}\langle \phi|. \end{aligned} \tag{13}$$

Acting on (12) with $\langle g, \phi| \equiv \langle g| \otimes \langle \phi|$ from the left, the scalar constraint becomes a functional partial differential equation of Wheeler–DeWitt type:

$$\left[\mathcal{C}_G\left(g, -\mathrm{i}\tfrac{\delta}{\delta g}\right) + \mathcal{C}_M\left(g, -\mathrm{i}\tfrac{\delta}{\delta g}, \phi, -\mathrm{i}\tfrac{\delta}{\delta \phi}\right)\right]\Psi[g,\phi] = 0, \tag{14}$$

where $\Psi[g,\phi] \equiv \langle g, \phi|\Psi\rangle$ is the wavefunctional of the combined gravity-matter system. We now try to look for a separable state, in the form $|\Psi\rangle = |\Psi_G\rangle \otimes |\Psi_M\rangle$, where $|\Psi_G\rangle \in \mathcal{H}_G$ and $|\Psi_M\rangle \in \mathcal{H}_M$, as a solution of this equation. Using the field representation (13), we write the wavefunctional $\Psi[g,\phi]$ as

$$\begin{aligned} \Psi[g,\phi] &\equiv \langle g, \phi|\Psi\rangle \\ &= \left(\langle g| \otimes \langle \phi|\right)\left(|\Psi_G\rangle \otimes |\Psi_M\rangle\right) \\ &= \langle g|\Psi_G\rangle\langle \phi|\Psi_M\rangle \\ &\equiv \Psi_G[g]\Psi_M[\phi]. \end{aligned} \tag{15}$$

Equation (14) can have separable solutions $\Psi[g,\phi] = \Psi_G[g]\Psi_M[\phi]$ if the functional differential operator $\mathcal{C}_M$ can be written as a product of two operators, denoted $\mathcal{K}_G$ and $\mathcal{K}_M$, depending only on $(g, \frac{\delta}{\delta g})$ and on $(\phi, \frac{\delta}{\delta \phi})$, respectively,

$$\mathcal{C}_M\left(g, -\mathrm{i}\tfrac{\delta}{\delta g}, \phi, -\mathrm{i}\tfrac{\delta}{\delta \phi}\right) = \mathcal{K}_G\left(g, \tfrac{\delta}{\delta g}\right)\mathcal{K}_M\left(\phi, \tfrac{\delta}{\delta \phi}\right). \tag{16}$$

If such operators $\mathcal{K}_G$ and $\mathcal{K}_M$ exist so that (16) holds, the scalar constraint equation (14) can be rewritten as

$$\Psi_M[\phi]\,\mathcal{C}_G\left(g, -\mathrm{i}\tfrac{\delta}{\delta g}\right)\Psi_G[g] = -\left[\mathcal{K}_G\left(g, \tfrac{\delta}{\delta g}\right)\Psi_G[g]\right]\left[\mathcal{K}_M\left(\phi, \tfrac{\delta}{\delta \phi}\right)\Psi_M[\phi]\right].$$

Dividing this with $\Psi_M[\phi]\,\mathcal{K}_G\left(g, \frac{\delta}{\delta g}\right)\Psi_G[g]$, assuming it is well-defined, we obtain

$$\frac{1}{\mathcal{K}_G\left(g, \frac{\delta}{\delta g}\right)\Psi_G[g]}\,\mathcal{C}_G\left(g, -\mathrm{i}\tfrac{\delta}{\delta g}\right)\Psi_G[g] = -\frac{1}{\Psi_M[\phi]}\,\mathcal{K}_M\left(\phi, \tfrac{\delta}{\delta \phi}\right)\Psi_M[\phi] = A,$$

where A is a constant, since the terms on the left and the right of the first equality depend on different sets of variables. Therefore, the above equation splits into two independent equations,

$$\begin{aligned} \left[\mathcal{C}_G\left(g, -\mathrm{i}\tfrac{\delta}{\delta g}\right) - A\,\mathcal{K}_G\left(g, \tfrac{\delta}{\delta g}\right)\right]\Psi_G[g] &= 0, \\ \left[\mathcal{K}_M\left(\phi, \tfrac{\delta}{\delta \phi}\right) + A\right]\Psi_M[\phi] &= 0, \end{aligned} \tag{17}$$

which are to be solved independently for $\Psi_G[g]$ and $\Psi_M[\phi]$, thus providing a separable solution of (14).

The whole procedure above rests on the assumption (16) that the matter part $\mathcal{C}_M$ of the scalar constraint operator can be written as a product of two operators $\mathcal{K}_G$ and $\mathcal{K}_M$. Our main argument is to demonstrate that the assumption (16) is never satisfied for the usual matter fields, due to the universal nature of the coupling of gravity to matter, ultimately dictated by the equivalence principle. Namely, given the structure of the classical scalar constraint for matter (8), the corresponding operator can be written as

$$\mathcal{C}^M(\hat{g},\hat{\pi}_g,\hat{\phi},\hat{\pi}_\phi) = \hat{\pi}_{\phi A}\hat{\nabla}_\perp{}^A{}_B\hat{\phi}^B - \frac{1}{N}\mathcal{L}_M(\hat{g},\hat{\pi}_g,\hat{\phi},\hat{\pi}_\phi), \tag{18}$$

where a certain ordering of the operators is assumed. The constraint (18) features the operator-valued matter Lagrangian $\mathcal{L}_M$. Therefore, in order to demonstrate that $\mathcal{C}_M$ does not satisfy the separability criterion (16) it is enough to demonstrate that the matter Lagrangian does not satisfy it. This can be done on a case-by-case basis, for each particular matter field. Invoking the equivalence principle, we can write the operator-valued Lagrangian for the scalar field coupled to gravity as

$$\mathcal{L}_M(\hat{g},\hat{\varphi},\partial\hat{\varphi}) = \frac{1}{2}\hat{e}\left[\hat{g}^{\mu\nu}(\partial_\mu\hat{\varphi})(\partial_\nu\hat{\varphi}) - m^2\hat{\varphi}^2 + U(\hat{\varphi})\right],$$

where $\hat{e}$ is the square-root of the minus determinant operator of the metric tensor,

$$\hat{e} \equiv \left[\frac{1}{4!}\varepsilon^{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma}\hat{g}_{\alpha\mu}\hat{g}_{\beta\nu}\hat{g}_{\gamma\rho}\hat{g}_{\delta\sigma}\right]^{\frac{1}{2}},$$

and U is some interaction potential of the scalar field. Ignoring the multiplicative factor $\hat{e}$ that acts only on $\mathcal{H}_G$, the Lagrangian is a sum of two types of terms: the kinetic term, containing the inverse metric $\hat{g}^{\mu\nu}$, and the mass and potential terms not featuring the gravitational field in any form. The sum cannot therefore be factored into the form $\mathcal{K}_G(\hat{g})\mathcal{K}_M(\hat{\phi},\partial\hat{\phi})$, since the Lagrangian is not a homogeneous function of the gravitational degrees of freedom. Even in the case of the massless free scalar field, i.e. when $m = 0$ and $U = 0$, the kinetic term is a sum of several different components of the metric and the derivatives of the scalar field,

$$\hat{g}^{00}(\partial_0\hat{\varphi})(\partial_0\hat{\varphi}) + \hat{g}^{01}(\partial_0\hat{\varphi})(\partial_1\hat{\varphi}) + \hat{g}^{12}(\partial_1\hat{\varphi})(\partial_2\hat{\varphi}) + \dots$$

and this still cannot be factored into a product of two operators $\mathcal{K}_G$ and $\mathcal{K}_M$.

In the case of the Dirac field, again invoking the equivalence principle, the operator-valued Lagrangian is given by (9),

$$\mathcal{L}_M(\hat{e},\hat{\omega},\hat{\psi},\hat{\bar{\psi}}) = \hat{e}\left(\frac{\mathrm{i}}{2}\hat{\bar{\psi}}\gamma^a\hat{e}^\mu{}_a\overset{\leftrightarrow}{\hat{\nabla}}_\mu\hat{\psi} - m\hat{\bar{\psi}}\hat{\psi}\right).$$

Like in the case of the scalar field, the kinetic and mass terms in the Lagrangian depend differently on the gravitational fields $\hat{e}^a{}_\mu$ and $\hat{\omega}^{ab}{}_\mu$, and $\mathcal{L}_M$ cannot be factored. Moreover, the kinetic term itself cannot be factored, since it is a sum of two terms (see equations (6)), only one of which contains the spin connection $\hat{\omega}^{ab}{}_\mu$.

Next, the operator-valued Lagrangian for the electromagnetic field coupled to gravity has the form

$$\mathcal{L}_M(\hat{g},\hat{A},\partial\hat{A}) = -\frac{1}{4}\hat{e}\,\hat{g}^{\mu\rho}\hat{g}^{\nu\sigma}\hat{F}_{\mu\nu}\hat{F}_{\rho\sigma},$$

where $\hat{F}_{\mu\nu} \equiv \partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu$. Applying the same argument as in the case of the free massless scalar field, this Lagrangian also cannot be factored into the form $\mathcal{K}_G \mathcal{K}_M$. The same argument also applies to the case of the non-Abelian Yang–Mills Lagrangians.

Summing up, given the ways the matter fields are coupled to gravity, based on the equivalence principle, we conclude that the separability criterion (16) is never satisfied for the physically relevant cases of scalar, spinor and vector fields. Therefore, according to the discussion above, the scalar constraint (12) should not admit separable state vectors into $\mathcal{H}_{\text{phys}}$.

Regarding the above analysis, it is important to emphasize the following. Namely, one should note that it is in principle possible for equation (14) to have product state solutions (15) despite the fact that it does not satisfy the separability criterion (16). In other words, the criterion (16) is a sufficient condition for the existence of product state solutions of (14), but it is not necessary, so its violation does not strictly imply the absence of product state solutions. Nevertheless, given the arguably highly complex structure of equation (14)—meaning that it represents a nonlinear functional partial differential equation of at least second order in g and ϕ—it is natural to regard any potential product state solutions as completely accidental. Moreover, it is questionable if the boundary conditions required for such solutions correspond to any realistic physical situation in nature, i.e. they could be irrelevant for realistic physics. Due to all these arguments, the existence of product state solutions, in spite of the violation of the separability criterion (16), is in our opinion an extraordinary claim, and as such requires extraordinary evidence. In other words, the burden of proof is in fact with the statement that any product state solution exists, rather than the opposite. Consequently, product states (15) are generically not elements of $\mathcal{H}_{\text{phys}}$, and even if one can prove that there exist some product states which do happen to belong to $\mathcal{H}_{\text{phys}}$, such states would arguably be completely accidental, with questionable relevance for physics. Otherwise, if there exists a whole class of separable states which solve (14) despite the violation of the criterion (16), there must be some deep eluding property of the scalar constraint equation, which is both completely unknown and very interesting to study.

Finally, while it turns out that the analysis of the scalar constraint equation (12) is sufficient for our conclusions, let us briefly mention the status of the remaining two constraint equations (11). First, the Gauss constraint $\hat{\mathcal{C}}_{ab}$ obviously admits separable state vectors. On the other hand, the situation with the 3-diffeo constraint $\hat{\mathcal{C}}_i$ is more complicated, and the conclusion depends on the type of the field. For example, in the case of the scalar field, from (5) we read that $\hat{\mathcal{C}}_i^M$ depends only on the scalar field and its momentum, which means that the constraint equation does admit separable state vectors. However, in the case of the Dirac field, from (7) we read that $\hat{\mathcal{C}}_i^M$ depends on the spin connection in addition to the Dirac field, and this dependence is not homogeneous in the spin connection, see (6). Thus, the 3-diffeo constraint equation does not admit separable state vectors. However, the behaviour of the Gauss and 3-diffeo constraint equations is redundant for our argument, since the scalar constraint equation (12) already suppresses separable state vectors for all fields, due to the dynamical form of the coupling of matter to gravity. Therefore, our initial assumption of local Poincaré symmetry can be weakened to the localisation of its translational subgroup, while the generators of the local Lorentz subgroup are irrelevant for our argument.

3. Entanglement in the path integral framework

In the previous section we have discussed the gauge-protected entanglement within the framework of the canonical quantisation of the gravitational field with matter. In this section, we focus instead on the path integral framework of quantisation. We analyse the entanglement

on the example of the Hartle–Hawking state, which is known to satisfy all constraints of the theory. In the next section, we are going to apply the results of this section to the concrete case of Regge quantum gravity.

First, we discuss an entanglement criterion for the case of pure overall state of the gravity and matter fields. We begin with a brief recapitulation of basic results from the standard QM and quantum information theory. A pure bipartite state $|\Psi\rangle_{12} \in \mathcal{H}_1 \otimes \mathcal{H}_2$ of systems 1 and 2 can be written in the Schmidt bi-orthogonal form (see, for example [20]):

$$|\Psi\rangle_{12} = \sum_i \sqrt{r_i}|\alpha_i\rangle_1 \otimes |\beta_i\rangle_2, \tag{19}$$

where $\{|\alpha_i\rangle_1\}$ and $\{|\beta_i\rangle_2\}$ are two sets of mutually orthogonal states from $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. The partial sub-system states are then given as

$$\hat{\rho}_1 = \sum_i r_i|\alpha_i\rangle_1 \otimes \langle\alpha_i|_1, \tag{20}$$

for the system 1, and analogously for the system 2. Squaring $\hat{\rho}_1$, we have

$$\hat{\rho}_1^2 = \sum_i r_i^2|\alpha_i\rangle_1 \otimes \langle\alpha_i|_1. \tag{21}$$

If the overall state $|\Psi\rangle_{12}$ is separable (i.e. a simple product state), the above sum in (20) will be trivial, consisting of a single projector onto the ray $|\alpha_1\rangle_1 \otimes \langle\alpha_1|_1$, with $r_1 = 1$. Thus, we have that $\hat{\rho}_1^2 = \hat{\rho}_1$, or simply, $\mathrm{Tr}\,\hat{\rho}_1^2 = \mathrm{Tr}\,\hat{\rho}_1 = 1$. In case the state $|\Psi\rangle_{12}$ is entangled, the sum (20) will consist of more than just one term, resulting in $(\forall\ i)\ r_i < 1$. Therefore, $(\forall\ i)\ r_i^2 < r_i$, and we finally have

$$\mathrm{Tr}\,\hat{\rho}_1^2 = \sum_i r_i^2 < \sum_i r_i = \mathrm{Tr}\,\hat{\rho}_1 = 1. \tag{22}$$

Due to the symmetry of the Schmidt form (19), the same is valid for the system 2 (for the formal proof of the above entanglement criterion (22), see for example [20]).

After this recapitulation of the standard results from QM, we proceed with the analysis of the bipartite system of the gravity (G) and matter (M) fields, applying the above entanglement criterion (22) to the case of quantum fields. For simplicity, we omit the subscripts G and M for pure states of gravity and matter, respectively.

Let $\mathcal{H}_{\mathrm{kin}} = \mathcal{H}_G \otimes \mathcal{H}_M$ be the combined kinematical gravity-matter Hilbert space. Denote the bases in $\mathcal{H}_G$ and $\mathcal{H}_M$ as $\{|g\rangle\}$ and $\{|\phi\rangle\}$, respectively. These are the eigenbases of the corresponding quantum field operators $\hat{g}$ and $\hat{\phi}$, evaluated on the $3D$ boundary $\Sigma_3 = \partial\mathcal{M}_4$ of the $4D$ spacetime manifold $\mathcal{M}_4$. The general state vector $|\Psi\rangle \in \mathcal{H}_{\mathrm{kin}}$ of the gravity-matter system can then be written as

$$|\Psi\rangle = \int \mathcal{D}g \int \mathcal{D}\phi\, \Psi[g,\phi]\, |g\rangle \otimes |\phi\rangle, \tag{23}$$

where $\Psi[g,\phi] = \langle g,\phi|\Psi\rangle$ is called the wavefunctional (in analogy to wavefunction from quantum mechanics), and the functional integrals over gravitational degrees of freedom g and matter degrees of freedom ϕ are assumed to be well defined in some way (in section 4 we present an explicit example of this). The bases $\{|g\rangle\}$ and $\{|\phi\rangle\}$ are assumed to be orthonormal, satisfying

$$\langle g|g'\rangle = \delta[g-g'], \qquad \langle\phi|\phi'\rangle = \delta[\phi-\phi'], \tag{24}$$

where the Dirac delta functional is assumed to satisfy the formal functional integral identities

$$\int \mathcal{D}g\, F[g]\, \delta[g-g'] \quad = \quad F[g'],$$
$$\int \mathcal{D}\phi\, F[\phi]\, \delta[\phi-\phi'] \quad = \quad F[\phi'], \tag{25}$$

for any functionals $F[g]$ and $F[\phi]$ belonging to some suitable relevant class.

From the state (23) one can construct a reduced density matrix $\hat{\rho}_M$ for matter fields, by taking the partial trace over gravitational degrees of freedom of the full density matrix $\hat{\rho} \equiv |\Psi\rangle \otimes \langle\Psi|$, as

$$\hat{\rho}_M = \mathrm{Tr}_G\, \hat{\rho} = \int \mathcal{D}g\, \langle g| \Big(|\Psi\rangle \otimes \langle\Psi| \Big) |g\rangle.$$

Substituting (23) we get

$$\hat{\rho}_M \quad = \quad \int \mathcal{D}g \int \mathcal{D}g' \int \mathcal{D}\phi' \int \mathcal{D}g'' \int \mathcal{D}\phi'' \\ \Psi^*[g',\phi']\, \Psi[g'',\phi'']\, \langle g| \Big(|g''\rangle \otimes |\phi''\rangle \otimes \langle g'| \otimes \langle\phi'| \Big) |g\rangle.$$

Using (24) and (25), the expression for the reduced density matrix can be evaluated to

$$\hat{\rho}_M = \int \mathcal{D}g \int \mathcal{D}\phi' \int \mathcal{D}\phi''\, \Psi^*[g,\phi']\Psi[g,\phi'']\, |\phi''\rangle \otimes \langle\phi'|. \tag{26}$$

Taking the square and using (24) and (25) again, one obtains

$$\hat{\rho}_M^2 \quad = \quad \int \mathcal{D}g \int \mathcal{D}g' \int \mathcal{D}\phi' \int \mathcal{D}\phi'' \int \mathcal{D}\phi''' \\ \Psi^*[g,\phi']\, \Psi[g,\phi'']\, \Psi^*[g',\phi''']\, \Psi[g',\phi']\, |\phi''\rangle \otimes \langle\phi'''|.$$

Finally, taking the trace over matter fields,

$$\mathrm{Tr}_M\, \hat{\rho}_M^2 = \int \mathcal{D}\phi\, \langle\phi|\hat{\rho}_M^2|\phi\rangle,$$

we get

$$\mathrm{Tr}_M\, \hat{\rho}_M^2 = \int \mathcal{D}g \int \mathcal{D}g' \int \mathcal{D}\phi \int \mathcal{D}\phi'\ \Psi^*[g,\phi']\, \Psi[g,\phi]\, \Psi^*[g',\phi]\, \Psi[g',\phi']. \tag{27}$$

Now we want to evaluate (27) for one specific state, namely the Hartle–Hawking state, denoted $|\Psi_{\mathrm{HH}}\rangle$. This state is known to satisfy the scalar constraint equation (12), see [14], and thus belongs to the physical Hilbert space $\mathcal{H}_{\mathrm{phys}}$. Our aim is to demonstrate that the Hartle–Hawking state is nonseparable, and the strategy is to argue that $\mathrm{Tr}_M\, \hat{\rho}_M^2 < 1$ for $\hat{\rho} = |\Psi_{\mathrm{HH}}\rangle \otimes \langle\Psi_{\mathrm{HH}}|$. The Hartle–Hawking state is defined by specifying the wavefunctional $\Psi[g,\phi]$ in (23) as

$$\Psi_{\mathrm{HH}}[g,\phi] = \mathcal{N} \int \mathcal{D}G \int \mathcal{D}\Phi\, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[g,\phi,G,\Phi]}. \tag{28}$$

Here $\mathcal{N}$ is a normalisation constant, the variables G and Φ (denoted with the capital letters) live in the bulk spacetime $\mathcal{M}_4$, while g and ϕ (denoted with lowercase letters) live on the boundary $\Sigma_3 = \partial\mathcal{M}_4$, as before. The path integrals are taken over the bulk while keeping the boundary fields constant. Finally, the total action functional S_{tot} has the following structure

$$S_{\text{tot}}[g,\phi,G,\Phi] = S_G[g,G] + S_M[g,\phi,G,\Phi], \tag{29}$$

where S_G is the action for the gravitational field (for example the Einstein–Hilbert action with a cosmological constant), while S_M is the action for the matter fields coupled to gravity—hence its dependence on both the gravitational and matter fields. See [14] for details on the construction of the expression (28).

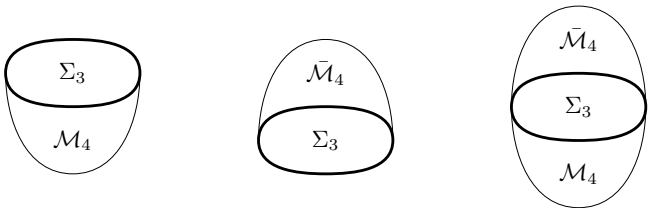
In order to analyse the expression (27) more efficiently, it is convenient to introduce the following quantity,

$$Z[\phi,\phi'] \equiv \int \mathcal{D}g\, \Psi_{\text{HH}}[g,\phi]\Psi^*_{\text{HH}}[g,\phi'], \tag{30}$$

which represents the matrix element of the reduced density matrix $\hat{\rho}_M$. Namely, by evaluating (26) for the Hartle–Hawking state, one obtains

$$\hat{\rho}_M = \int \mathcal{D}\phi \int \mathcal{D}\phi'\, Z[\phi,\phi']\, |\phi\rangle \otimes \langle\phi'|. \tag{31}$$

In addition, $Z[\phi,\phi']$ has an important geometric structure. Namely, one can consider two copies of the spacetime manifold $\mathcal{M}_4$, where the boundary Σ_3 of the first copy features the fields g,ϕ, while the boundary of the second copy features the fields g,ϕ', i.e. such that the gravitational field g is the same, while matter fields ϕ and ϕ' are different on the boundaries. Then one takes the second copy of $\mathcal{M}_4$, inverts it with respect to the boundary Σ_3 (the result is denoted as $\bar{\mathcal{M}}_4$), and glues it to the first copy along the common boundary, to obtain a manifold $\mathcal{M}_4 \cup \bar{\mathcal{M}}_4$, which has no boundary. This can be illustrated by the following diagrams:



The quantity $Z[\phi,\phi']$ is then obtained by integrating over all gravitational degrees of freedom, and all bulk matter degrees of freedom, weighted by the kernel $\mathrm{e}^{\mathrm{i}S_{\text{tot}}}$ of the Hartle–Hawking wavefunction (28). This construction is important because the trace of $Z[\phi,\phi']$ is the state sum of the gravitational and matter fields over the manifold $\mathcal{M}_4 \cup \bar{\mathcal{M}}_4$:

$$\int \mathcal{D}\phi\, Z[\phi,\phi] = Z \equiv \int \mathcal{D}G \int \mathcal{D}\Phi\, \mathrm{e}^{\mathrm{i}S[G,\Phi]}. \tag{32}$$

Here, $S[G,\Phi]$ is the total gravity-matter action similar to (29), defined over the manifold $\mathcal{M}_4 \cup \bar{\mathcal{M}}_4$, and thus features no boundary fields. From (30)–(32) it is then easy to see that the normalization of the state sum, $Z = 1$, and simultaneously the normalization of the reduced density matrix, $\mathrm{Tr}\,\hat{\rho}_M = 1$, i.e.

$$\int \mathcal{D}\phi\, Z[\phi,\phi] = 1, \tag{33}$$

are equivalent to the normalisation of the Hartle–Hawking state, $\langle\Psi_{\text{HH}}|\Psi_{\text{HH}}\rangle = 1$. Finally, from the definition (30) it is easy to see that $Z[\phi,\phi']$ is self-adjoint,

$$Z[\phi,\phi'] = Z^*[\phi',\phi],$$

as the matrix elements of the density matrix $\hat{\rho}_M$ are supposed to be.

Returning to the evaluation of (27) for the Hartle–Hawking state, one can use (30) to rewrite it into the compact form

$$\mathrm{Tr}_M\, \hat{\rho}_M^2 = \int \mathcal{D}\phi \int \mathcal{D}\phi' \, \left|Z[\phi, \phi']\right|^2 . \tag{34}$$

At this point the general analysis cannot proceed any further, since the right-hand side cannot be evaluated explicitly without specifying the details of the theory. The calculation will therefore proceed further in the next section, where we consider one detailed model of quantum gravity with matter.

Despite the inability to evaluate the integral (34) in the general case, one can give a qualitative argument that the result is not equal to one, leading to the nonseparability of the Hartle–Hawking state. Namely, given the definition (28) of the Hartle–Hawking state, it is easy to see that it essentially depends on two quantities—the normalisation constant $\mathcal{N}$, and the choice of the action S_{tot}. The normalisation constant is fixed by the requirement that (33) holds. This leaves the value of the integral (34) depending solely on the choice of the classical action of the theory. It is qualitatively straightforward to see that different choices of the action will lead to different values of $\mathrm{Tr}_M\, \hat{\rho}_M^2$, so any generic choice of S_{tot} is likely to give $\mathrm{Tr}_M\, \hat{\rho}_M^2 < 1$. A tentative choice for (29) would be the Einstein–Hilbert action for S_G and the Standard Model of elementary particle physics for S_M, based on the gauge group $SU(3) \times SU(2) \times U(1)$. However, we know that the Standard Model action is incomplete, for example due to the fact that dark matter is not included in the description. Therefore, the choice of the classical action is a sort of a moving target, and it is unlikely that any candidate action we choose will give $\mathrm{Tr}_M\, \hat{\rho}_M^2 = 1$. In this sense, one can only conclude that in a generic case the Hartle–Hawking state is nonseparable, supporting the abstract argument from section 2.

Finally, let us note that our assumption of local Poincaré gauge symmetry implies that we are discussing the Lorentzian path integral formulation of the theory. In contrast, within the Euclidean approach, the Hartle–Hawking state has some problematic characteristics, see for example [21] and references therein.

4. Regge quantum gravity example

In this section we will present a short review of the Regge quantum gravity model coupled to scalar matter, and then use this model to evaluate (34) for the Hartle–Hawking state. The Regge quantum gravity model is intimately connected to the covariant loop quantum gravity research framework [22, 23], its generalisations [24–26], and various related research areas [27, 28] (see also [29] for an interesting connection to the noncommutative geometry approach in the $3D$ case). Nevertheless, it can be introduced and studied as a simple standalone model of quantum gravity independent of any other context, as was done in [30], where some preliminary results regarding the entanglement in the Hartle–Hawking state have been announced.

4.1. Formalism of Regge quantum gravity

The Regge quantum gravity model is arguably the simplest toy-model of quantum gravity constructed by providing a rigorous definition for the gravitational path integral, generically denoted as

$$Z_G = \mathcal{N} \int \mathcal{D}g \, \mathrm{e}^{\mathrm{i}S_{\mathrm{EH}}[g]}, \tag{35}$$

where $S_{\mathrm{EH}}[g]$ is the Einstein–Hilbert action for general relativity. The construction of the path integral follows Feynman's original idea of path integral definition-by-discretisation. We begin by passing from a smooth $4D$ spacetime manifold $\mathcal{M}_4$ to a piecewise-linear $4D$ manifold, most commonly a triangulation $T(\mathcal{M}_4)$. This structure naturally features 4-simplices σ as basic building blocks, which themselves consist of tetrahedra τ, triangles Δ, edges ϵ and vertices v. The invariant quantities associated to these objects are the 4-volume of the 4-simplex ${}^{(4)}V_\sigma$, the 3-volume of the tetrahedron ${}^{(3)}V_\tau$, the area of the triangle A_Δ and the length of the edge l_ϵ, respectively, while the vertices do not have nontrivial quantities assigned to them.

It is important to emphasise that the edge lengths are most fundamental of all these quantities, since one can always uniquely express ${}^{(4)}V_\sigma$, ${}^{(3)}V_\tau$ and A_Δ as functions of l_ϵ. For example, the most well-known is the Heron formula for the area of a triangle in terms of its three edge lengths,

$$A_\Delta(l) = \sqrt{s(s-l_1)(s-l_2)(s-l_3)}, \qquad s \equiv \frac{l_1+l_2+l_3}{2},$$

where the three edges $\epsilon = 1,2,3$ belong to the triangle Δ.

Given a spacetime triangulation, the Einstein–Hilbert action of general relativity,

$$S_{\mathrm{EH}}[g] = -\frac{1}{16\pi l_p^2}\int_{\mathcal{M}_4} \mathrm{d}^4x\sqrt{-g}\,R(g),$$

can be reformulated in terms of edge lengths of the triangulation as the Regge action

$$S_R[l] = -\frac{1}{8\pi l_p^2}\sum_{\Delta\in T(\mathcal{M}_4)} A_\Delta(l)\delta_\Delta(l),$$

where δ_Δ is the so-called deficit angle at triangle Δ, measuring the amount of spacetime curvature around Δ. See [31] and [27] for details and a review.

Once the classical action for general relativity has been adapted to a piecewise-linear manifold structure, we can take the edge lengths of the edges in the triangulation as the fundamental degrees of freedom of the theory, and define the gravitational path integral (35) as:

$$Z_G = \mathcal{N}\int_D \prod_{\epsilon\in T(\mathcal{M}_4)} \mathrm{d}l_\epsilon\,\mu(l)\mathrm{e}^{\mathrm{i}S_R[l]}. \tag{36}$$

Here $\mathcal{N}$ is a normalisation constant, while $\mu(l)$ is the measure term which ensures the convergence of the state sum Z_G. For the purpose of this paper, we choose the exponential measure

$$\mu(l) = \exp\left(-\frac{1}{L_\mu^4}\sum_{\sigma\in T(\mathcal{M}_4)} {}^{(4)}V_\sigma(l)\right), \tag{37}$$

where $L_\mu > 0$ is a constant and a free parameter of the model (see [32–34] for motivation and analysis). Note that the sum of the 4-volumes of all 4-simplices gives the total 4-volume of the triangulation $T(\mathcal{M}_4)$, and will sometimes be denoted simply as V_4.

The choice of edge lengths as the fundamental gravitational degrees of freedom in (36) determines the integration domain D as a subset of the Cartesian product $(\mathbb{R}_0^+)^E$, where E is the total number of edges in $T(\mathcal{M}_4)$, while $\mathbb{R}_0^+$ is the maximum integration domain of each individual edge length. We should note that D is a strict subset of $(\mathbb{R}_0^+)^E$ due to the presence of triangle inequalities which must be satisfied for all triangles, tetrahedra and 4-simplices in a given triangulation.

Once we have defined the gravitational path integral (35) via the state sum (36), it is straightforward to generalise this definition to the situation which includes matter fields. For simplicity, we will discuss only a single real scalar field φ, although it is not a problem to include other fields as well. The path integral we are interested in can be denoted as

$$Z_{G+M} = \mathcal{N} \int \mathcal{D}g \int \mathcal{D}\varphi \, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[g,\varphi]}, \tag{38}$$

where $S_{\mathrm{tot}}[g,\varphi]$ is the sum of the Einstein–Hilbert action and the action for the scalar field in curved spacetime,

$$\begin{aligned} S_{\mathrm{tot}}[g,\varphi] &= -\frac{1}{16\pi l_p^2} \int_{\mathcal{M}_4} \mathrm{d}^4x\sqrt{-g}\, R(g) \\ &\quad + \frac{1}{2} \int_{\mathcal{M}_4} \mathrm{d}^4x\sqrt{-g}\, \left[g^{\mu\nu}(\partial_\mu\varphi)(\partial_\nu\varphi) + m^2\varphi^2 + U(\varphi) \right], \end{aligned}$$

where $U(\varphi)$ is a self-interaction potential of the scalar field. The corresponding lattice version of this action is given as

$$\begin{aligned} S_{\mathrm{tot}}[l,\varphi] &= -\frac{1}{8\pi l_p^2} \sum_{\Delta\in T(\mathcal{M}_4)} A_\Delta(l)\delta_\Delta(l) + \\ &\quad + \frac{1}{2} \sum_{\sigma\in T(\mathcal{M}_4)} {}^{(4)}V_\sigma(l) g^{\mu\nu}_{(\sigma)}(l) \partial\varphi_\mu \partial\varphi_\nu \\ &\quad + \frac{1}{2} \sum_{v\in T(\mathcal{M}_4)} {}^{(4)}V^*_v(l) \left[m^2\varphi_v^2 + U(\varphi_v) \right]. \end{aligned} \tag{39}$$

Here, a value of the scalar field $\varphi_v \in \mathbb{R}$ is assigned to each vertex $v \in T(\mathcal{M}_4)$. Given any 4-simplex $\sigma \in T(\mathcal{M}_4)$, one can label its five vertices as $0,1,2,3,4$, and then define a skew-coordinate system taking the vertex 4 as the origin and edges $4-0$, $4-1$, $4-2$, $4-3$, respectively as coordinate lines for coordinates x^μ, $\mu \in \{0,1,2,3\}$. In these coordinates, the derivative $\partial_\mu\varphi$ is replaced by the finite difference between the values of the field at the vertex $v=\mu$ and at the coordinate origin of the 4-simplex σ (divided by the distance between them),

$$\partial\varphi_\mu \equiv \frac{\varphi_\mu - \varphi_4}{l_{\mu 4}}.$$

In addition, the metric tensor between vertices μ and ν is given in terms of edge lengths as

$$g^{(\sigma)}_{\mu\nu}(l) \equiv \frac{l^2_{\mu 4} + l^2_{\nu 4} - l^2_{\mu\nu}}{2 l_{\mu 4} l_{\nu 4}},$$

while $g^{\mu\nu}_{(\sigma)}(l)$ is its inverse matrix. Finally, ${}^{(4)}V^*_v(l)$ is the 4-volume of the 4-cell surrounding the vertex v in the Poincaré dual lattice of the triangulation $T(\mathcal{M}_4)$.

After we have defined the classical action on $T(\mathcal{M}_4)$, we finally proceed to define the path integral (38) as the state sum:

$$Z_{G+M} = \mathcal{N} \int \prod_{\epsilon\in T(\mathcal{M}_4)} \mathrm{d}l_\epsilon \, \mu(l) \int \prod_{v\in T(\mathcal{M}_4)} \mathrm{d}\varphi_v \, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[l,\varphi]}. \tag{40}$$

Here, the domain of integration for the scalar field is the Cartesian product $\mathbb{R}^V$, where V is the total number of vertices in the triangulation.

The state sum (40) defines one concrete QG model, called the Regge quantum gravity model. While it goes without saying that this is just a toy model, it is nevertheless a realistic one, since it is finite and has a correct semiclassical continuum limit (see [32] for proofs). Therefore it can be used to study various aspects of quantum gravity, including the entanglement between gravity and matter fields, as we discuss next.

4.2. Calculation of the trace formula

Having formulated the Regge quantum gravity model and having the state sum (40) in hand, we can proceed to study the entanglement between gravity and matter, in particular by evaluating the expression for the trace of $\hat{\rho}_M^2$ given by equation (34). In order to evaluate it, we first need to formulate the Hartle–Hawking state (28) in the framework of Regge quantum gravity model, then work out the matrix elements of the reduced density matrix (30), and finally plug them into (34) to obtain a number. If this number is different from 1, we can conclude that the Hartle–Hawking state features entanglement between the gravitational and matter fields.

We begin by formulating the Hartle–Hawking state (28). Consider a 4-manifold $\mathcal{M}_4$ with a nontrivial boundary $\Sigma_3 = \partial\mathcal{M}_4$, such that the triangulation $T(\mathcal{M}_4)$ induces a triangulation $T(\Sigma_3)$ on the boundary. In this sense we can distinguish the vertices, edges, areas, and tetrahedra which belong to the boundary triangulation $T(\Sigma_3)$ (from now on shortly called 'boundary', and denoted as ∂T), from the vertices, edges, areas, tetrahedra and 4-simplices belonging to $T(\mathcal{M}_4)$ but not to $T(\Sigma_3)$ (from now on shortly called 'bulk', and denoted as T). Since the Regge quantum gravity model encodes gravitational degrees of freedom as lengths of the edges, and matter degrees of freedom as real numbers attached to vertices, we can easily split them into boundary variables l_ϵ, φ_v and bulk variables L_ϵ, Φ_v, where we maintain our previous convention to denote the bulk variables with capital letters and boundary variables with lowercase letters.

Given the bulk and the boundary, we use the formulation of the Regge quantum gravity state sum (40) to write down the Hartle–Hawking wavefunction as

$$\Psi_{\mathrm{HH}}[l,\varphi] = \mathcal{N}\int \prod_{\epsilon\in T} \mathrm{d}L_\epsilon\, \mu(l,L) \int \prod_{v\in T} \mathrm{d}\Phi_v\, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[l,\varphi,L,\Phi]}. \tag{41}$$

Next we want to construct the matrix elements of the reduced density matrix (30). To this end, we need two copies of the Hartle–Hawking state: one with matter fields φ_v on the boundary ∂T of the bulk T, and the other with matter fields φ'_v on the boundary ∂T of the bulk $\bar{T}$ defined as the mirror-reflection of T with respect to the boundary ∂T. This mirror-reflection gives rise to an additional overall minus sign in the action (39) which is then cancelled by the complex conjugation of the imaginary unit in the exponent of the second Hartle–Hawking wavefunction in (30). Integrating over the boundary edge lengths, we end up with:

$$Z[\varphi,\varphi'] = |\mathcal{N}|^2 \int \prod_{\epsilon\in T\cup\bar{T}\cup\partial T} \mathrm{d}L_\epsilon\, \mu(L) \int \prod_{v\in T\cup\bar{T}} \mathrm{d}\Phi_v\, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[\varphi,\varphi',L,\Phi]}. \tag{42}$$

Note that all edge lengths are being integrated over in the 'total' triangulation $T\cup\bar{T}\cup\partial T$ (and we have thus denoted them all with a capital letter L for simplicity). In contrast, the scalar field is being integrated only over the two bulks $T\cup\bar{T}$, while the boundary scalar field values φ,φ' remain fixed on two identical copies of the boundary ∂T. Also, note that

$$S_{\mathrm{tot}}[\varphi,\varphi',L,\Phi] \equiv S_{\mathrm{tot}}[\varphi,L,\Phi]\Big|_{T\cup\partial T} + S_{\mathrm{tot}}[\varphi',L,\Phi]\Big|_{\bar{T}\cup\partial T},$$

where the boundary edge lengths l have been relabelled as L and reabsorbed into the set of bulk edge lengths.

The next step one should perform is to take the trace of (42) and equate it to 1 as in (33), in order to make sure that the Hartle–Hawking wavefunction (41) is properly normalised. This leads to the equation

$$|\mathcal{N}|^2 \int \prod_{\epsilon\in T\cup\bar{T}\cup\partial T} \mathrm{d}L_\epsilon\, \mu(L) \int \prod_{v\in T\cup\bar{T}\cup\partial T} \mathrm{d}\Phi_v\, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[L,\Phi]} = 1,$$

which determines the normalisation constant $\mathcal{N}$ up to an overall phase factor. Note that the boundary scalar fields φ have been integrated over and consequently reabsorbed into the bulk variables Φ, similarly to edge lengths L. Both the integration over L and the integration over Φ is now being performed over the 'total' triangulation $T\cup\bar{T}\cup\partial T$ which has no boundary.

As the final step of the construction of the trace formula (34), we substitute (42) and $\mathcal{N}$ into it, to obtain:

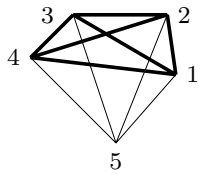
$$\mathrm{Tr}_M\,\hat{\rho}_M^2 = \frac{\displaystyle\int \prod_{v\in\partial T} \mathrm{d}\varphi_v \int \prod_{v\in\partial T} \mathrm{d}\varphi'_v \left| \int \prod_{\epsilon\in T\cup\bar{T}\cup\partial T} \mathrm{d}L_\epsilon\, \mu(L) \int \prod_{v\in T\cup\bar{T}} \mathrm{d}\Phi_v\, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[\varphi,\varphi',L,\Phi]} \right|^2}{\left(\displaystyle\int \prod_{\epsilon\in T\cup\bar{T}\cup\partial T} \mathrm{d}L_\epsilon\, \mu(L) \int \prod_{v\in T\cup\bar{T}\cup\partial T} \mathrm{d}\Phi_v\, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[L,\Phi]} \right)^2}. \tag{43}$$

This is the final expression we set out to derive. It represents a concrete realisation of the trace formula (34), it is completely well defined, and can in principle be evaluated. In practice, though, for a generic choice of the triangulation, this expression is very hard to evaluate even numerically. Therefore, in what follows we shall enforce some very hard approximations in order to make it more manageable for study. Nevertheless, by looking at the structure of the numerator and the denominator, one can already see that the two expressions can be equal to each other only in some very special cases, if at all. However, the dependence of the action S_{tot} on the boundary and bulk variables is such that one cannot rely on any special mathematical properties of the action which could help make the final result be 1, for a generic choice of the spacetime triangulation. In this sense, we can conjecture already at this level that in generic cases we have

$$\mathrm{Tr}_M\,\hat{\rho}_M^2 < 1,$$

as we wanted to demonstrate.

But in order to give a more convincing argument, let us study a special case and try to evaluate this trace to the very end. The simplest possible example of a triangulation T is a single 4-simplex. Labelling its vertices as $1, 2, 3, 4, 5$, we can depict it with a following diagram:



The 4-simplex has five boundary tetrahedra, namely

$$\tau_{1234}, \quad \tau_{1235}, \quad \tau_{1245}, \quad \tau_{1345}, \quad \tau_{2345}.$$

The first tetrahedron, τ_{1234}, is depicted with thick edges, and we will choose it to be the boundary ∂T. Since we do not want the four remaining tetrahedra to belong to the boundary, we will glue them onto each other in pairs, as

$$\tau_{1235} \equiv \tau_{1245}, \qquad \tau_{1345} \equiv \tau_{2345}.$$

This means that every point in τ_{1235} is identified with the corresponding point in τ_{1245}, and similarly with the other pair of tetrahedra. In this way we obtain a manifold with a nontrivial topology, but described with only five vertices and one boundary tetrahedron. In order for this gluing to be consistent, the gravitational and matter degrees of freedom living on $T \cup \partial T$ must satisfy the following constraints:

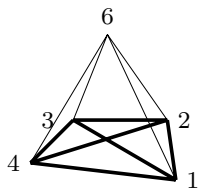
$$\begin{aligned} l_{14} = l_{23} = l_{24} = l_{13}, \qquad L_{25} = L_{15}, \qquad L_{45} = L_{35}, \\ \varphi_2 = \varphi_1, \qquad \varphi_4 = \varphi_3. \end{aligned} \tag{44}$$

This leaves us with the following independent degrees of freedom living on the 4-simplex:

$$l_{12}, \qquad l_{13}, \qquad L_{15}, \qquad l_{34}, \qquad L_{35}, \qquad \varphi_1, \qquad \varphi_3, \qquad \Phi_5,$$

where we have denoted the bulk degrees of freedom with capital letters and boundary degrees of freedom with lowercase letters. The 4-simplex diagram above is the graphical representation of the Hartle–Hawking wavefunction $\Psi_{\mathrm{HH}}[l, \varphi]$ (41).

Next we construct $\bar{T}$. Since the boundary tetrahedron ∂T defines a single 3-dimensional hypersurface, there is precisely one axis in 4-dimensional space which is orthogonal to ∂T. Performing the reflection of T with respect to ∂T is therefore identical to reversing the orientation of this orthogonal axis. In this way we construct another 4-simplex, with vertices labeled $1, 2, 3, 4, 6$ and depicted as



One can see that the main difference between the 4-simplex σ_{12346} and the previously constructed 4-simplex σ_{12345} is that the vertex 6 is on the 'opposite side' of the tetrahedron τ_{1234} as compared to the vertex 5 of σ_{12345}.

Like we did for σ_{12345}, we again want to glue the boundary tetrahedra pairwise, so that only the tetrahedron τ_{1234} remains as the boundary $\partial \bar{T}$. The pairwise gluing of tetrahedra

$$\tau_{1236} \equiv \tau_{1246}, \qquad \tau_{1346} \equiv \tau_{2346}$$

gives rise to the constraints

$$\begin{aligned} l_{14} = l_{23} = l_{24} = l_{13}, \qquad L_{26} = L_{16}, \qquad L_{46} = L_{36}, \\ \varphi'_2 = \varphi'_1, \qquad \varphi'_4 = \varphi'_3, \end{aligned}$$

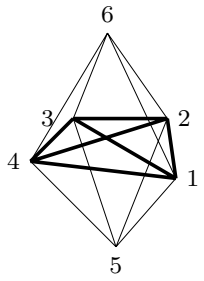
where only the constraints containing the vertex 6 are additional to (44). This leaves us with the following independent degrees of freedom living on σ_{12346}:

$$l_{12}, \qquad l_{13}, \qquad L_{16}, \qquad l_{34}, \qquad L_{36}, \qquad \varphi'_1, \qquad \varphi'_3, \qquad \Phi_6.$$

As noted in the general discussion leading to equation (43), the matter degrees of freedom on the boundary of T are different than the corresponding degrees of freedom living on the boundary of $\bar{T}$, despite the fact that the boundary is identical, $\partial \bar{T} \equiv \partial T$. To that end, we have added a prime to φ in the above equations. Like for the 4-simplex σ_{12345}, the diagram of the

4-simplex σ_{12346} above is the graphical representation of the (complex-conjugate) Hartle–Hawking wavefunction $\Psi^*_{\mathrm{HH}}[l,\varphi']$.

At this point we are ready to glue T and $\bar{T}$ along the common boundary ∂T, to obtain the manifold $T \cup \bar{T} \cup \partial T$ which has no boundary. It is depicted on the diagram below.



It consists of two 4-simplices σ_{12345} and σ_{12346} constructed above and glued along the common tetrahedron τ_{1234}. The full set of independent gravitational degrees of freedom is

$$l_{12}, \qquad l_{13}, \qquad l_{34}, \qquad L_{15}, \qquad L_{16}, \qquad L_{35}, \qquad L_{36},$$

while the independent matter degrees of freedom are

$$\varphi_1, \qquad \varphi_3, \qquad \varphi'_1, \qquad \varphi'_3, \qquad \Phi_5, \qquad \Phi_6.$$

This diagram is the graphical representation for the matrix element $Z[\varphi,\varphi']$ of the reduced density matrix $\hat{\rho}_M$ (see equations (42) and (30)).

Applying the general trace formula (43) to our case then gives

$$\mathrm{Tr}_M\, \hat{\rho}_M^2 = \frac{\displaystyle\int \mathrm{d}\varphi_1 \mathrm{d}\varphi_3 \mathrm{d}\varphi'_1 \mathrm{d}\varphi'_3 \left| \int \mathrm{d}^7 L\, \mu(L) \int \mathrm{d}\Phi_5 \mathrm{d}\Phi_6\, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[\varphi,\varphi',L,\Phi]} \right|^2}{\left(\displaystyle\int \mathrm{d}^7 L\, \mu(L) \int \mathrm{d}^4\Phi\, \mathrm{e}^{\mathrm{i}S_{\mathrm{tot}}[L,\Phi]} \right)^2}, \tag{45}$$

where

$$\mathrm{d}^7 L \equiv \mathrm{d}l_{12}\mathrm{d}l_{13}\mathrm{d}l_{34}\mathrm{d}L_{15}\mathrm{d}L_{16}\mathrm{d}L_{35}\mathrm{d}L_{36},$$

and

$$\mathrm{d}^4\Phi \equiv \mathrm{d}\varphi_1 \mathrm{d}\varphi_3 \mathrm{d}\Phi_5 \mathrm{d}\Phi_6.$$

Note that the action in the denominator is evaluated using $\varphi'_1 = \varphi_1$ and $\varphi'_3 = \varphi_3$, as explained in the general discussion above. In order to make the equation (45) fully explicit, we need to choose the values of the free parameters in the classical action (39) and the measure (37). The parameters of the action are the Planck length l_p, the mass m of the scalar field, and the self-interaction potential $U(\varphi)$. For the purpose of this example, the simplest possible choice is the free massless scalar field, so that we have

$$l_p = 10^{-35}\ \mathrm{m}, \qquad m = 0, \qquad U(\varphi) = 0.$$

Second, the measure (37) contains a single free parameter L_μ. This parameter can be connected to the value of the effective cosmological constant Λ, via the relation

$$\Lambda = \frac{l_p^2}{2L_\mu^4},$$

see [32–34] for details. Taking the observed value $\Lambda = 10^{-52}$ m^{-2} (also often quoted as a dimensionless product $\Lambda l_p^2 = 10^{-122}$), we obtain

$$L_\mu = 10^{-5} \text{ m}.$$

Using these numeric values of the parameters, the right-hand side of (45) is fully specified, and can be evaluated using a computer. However, in order to render the calculation more manageable, for the purpose of this paper we instead choose to evaluate (45) with $L_\mu = 10^{-33}$ m, which corresponds to a larger cosmological constant, $\Lambda l_p^2 = 10^{-8}$, to speed up the convergence of the Monte-Carlo integration method. The result is strictly less than one,

$$\text{Tr}_M\, \hat{\rho}_M^2 = 0.977 \pm 0.002,$$

as we had set out to demonstrate. Note that, although close to one, the above result is: (i) strictly smaller than one (within the computational error); (ii) obtained within extremely simplified toy model whose system consists of only two 4-simplices of spacetime. Thus, our result can serve as a proof of principle that gravity-matter entanglement is always present. The total amount of such entanglement in realistic models, as well as its spatial distribution, remains to be further explored. Namely, note that even though the approximation of product gravity-matter states has been up to now successfully applied, the *overall* entanglement between the two systems, considered within complex realistic situations/models, does not at all have to be small, nor its effects negligible. Indeed, the standard entanglement that is considered to cause the decoherence of matter by the environment and the quantum-to-classical transition has profoundly striking effects, despite the fact of being difficult to characterise, evaluate and manipulate.

5. Conclusions

5.1. Summary of the results

We analyse the quantum gravity coupled to the most common matter fields (namely, scalar, spinor and vector fields), and show that the gravity and matter are generically entangled, as a consequence of the nonseparability of the scalar constraint $\mathcal{C}$, and in some cases the 3-diffeo constraint $\mathcal{C}_i^M$. Thus, simple separable gravity-matter product states are excluded from the physical Hilbert space, unless the constraint equations feature some deep unknown property which allows for the invariance of a whole class of product states. We demonstrate this in two different ways: (i) within the general abstract nonperturbative canonical formalism, by directly analysing the mathematical structure of the constraints, and (ii) within the path integral formalism, by directly checking for entanglement of the Hartle–Hawking state in the Regge model of quantum gravity.

5.2. Discussion of the results

This *gauge-protected* decoherence due to the entanglement (in contrast to the standard *'for all practical purposes'* dynamical one) offers a possibly deeper fundamental explanation of the long-standing problem of the quantum-to-classical transition: the matter does not *decohere*, it is by default *decohered*.

Any potential entanglement, either dynamical or gauge-protected one, depends on the details of the coupling between matter and gravity. For the purpose of this paper, the coupling is prescribed by the strong equivalence principle, which states that the equations of motion for

all matter fields must locally be identical to the equations of motion for those fields in flat spacetime. This is implemented by choosing the action for matter fields with minimal coupling prescription, and employed in both the canonical and the path integral frameworks. We should stress that the validity of the strong equivalence principle is a sufficient, but potentially not a necessary assumption for our main result. Namely, it is plausible that nonminimal coupling choices, involving explicit spacetime curvature terms in the matter Lagrangian, could also lead to the conclusion that entanglement between gravity and matter is unavoidable. However, it is also possible that one could come up with some particular complicated choice of nonminimal coupling which does admit some nonentagled states. In order to avoid complicating the analysis with such cases, given that nonminimal coupling between gravity and matter has absolutely no experimental evidence in its favor so far, we have chosen to assume the validity of the strong equivalence principle throughout the paper.

In standard QM entanglement is a generic consequence of the interaction. Nevertheless, there exist alternative mechanisms for creating it, such as the indistinguishability of identical particles, leading to effective 'exchange interactions'. This new gauge-protected gravity-matter entanglement can thus introduce additional 'effective interaction', which can possibly result in corrections to Einstein's weak equivalence principle (see for example [35]).

It is interesting to note that a possible peculiar impact of the quantised gravity to the whole decoherence programme was already inferred in Zurek's seminal paper [36], where on page 1520 the author writes: (the assumption of pairwise interactions) 'is customary and clear, even though it may prevent one from even an approximate treatment of the gravitational interaction beyond its Newtonian pairwise form'. Our result confirms Zurek's disclaimer—gravity (environment $\mathcal{E}$) is generically entangled with the *whole* matter (both the system $\mathcal{S}$ and the apparatus $\mathcal{A}$), that way allowing for non-trivial tripartite system-apparatus-environment *effective* interaction of the form $\mathcal{H}_{\mathcal{SAE}}$, explicitly excluded in [36]. In other words, the environment (spacetime) interaction with the matter could potentially disturb the system-apparatus correlations, thus violating the stability criterion of a faithful measurement (see [37], p 1271).

As a consequence of generic gravity-matter entanglement, the effective interaction between gravity and matter forbids the existence of a single background spacetime. Thus, when concerning quantum effects of gravity, one cannot talk of 'matter in a point of space', confirming the conjecture that spacetime is an 'emergent phenomenon'. In contrast to this, Penrose argues that spacetime, seen as a (four-dimensional) differentiable manifold, does not support superpositions of massive bodies and the corresponding (relative) states of gravity (i.e. the gravity-matter entanglement), leading to the objective collapse onto the product states of matter and (classical) spacetime [38]. Our result can therefore be treated as a possible criterion for a plausible candidate theory of quantum gravity.

Finally, not allowing product states between the matter and gravity is in tune with the relational approach to physics [22], in particular to quantum gravity (note that the original name for the many-world interpretation of QM was the *'Relative State' Formulation of Quantum Mechanics* [39]). See also [40] for an interesting treatment of relative state and decoherence approaches.

5.3. Relation to common quantum gravity research programs

In order to discuss our results in the context of various quantum gravity research programs, note that the gauge-protected entanglement between gravity and matter should exist in any model of quantum gravity with matter which respects local Poincaré symmetry. In this sense, various approaches to quantum gravity can be classified into four distinct categories.

(i) The first category represents models which explicitly respect (or at least aim to respect) local Poincaré symmetry. These include nonperturbative string theory/M-theory [41–43], loop quantum gravity [22, 23], Wheeler–DeWitt quantization [44, 45], and similar approaches.
(ii) The second category represents models in which local Poincaré symmetry is explicitly broken. These include perturbative quantum gravity [46], petrurbative string theory [43], causal dynamical triangulations approach [28], doubly-special relativity models [47], Hořava–Lifshitz gravity [48], various nonrelativistic quantization proposals, and so on.
(iii) The third category represents models in which it is not clear whether local Poincaré symmetry is broken or not. For example, in the asymptotic safety approach [49] this may depend on the properties of the fixed point. In noncommutative geometry [50, 51] it depends on the particular choice of the algebra. In higher-derivative theories and theories with propagating torsion [52] it may depend on various details of the model, etc.
(iv) Finally, the fourth category represents models which have not been developed enough to allow for coupling of matter fields. In models like entropic gravity [53, 54] and causal set theory [55, 56], it is not obvious how to couple matter fields to gravity, and whether this coupling would violate local Poincaré invariance or not.

It should be clear that our results apply to the first category of quantum gravity models, while for other three categories it either does not apply, or it is an open question. We should also state that the validity of local Poincaré symmetry is ultimately an experimental question, one over which various quantum gravity proposals may disagree.

In relation to the previous comment, it is worthwhile to also discuss the impact of possible anomalies to the gauge protected entanglement. As we have discussed in the final paragraph of section 2, the entanglement is a consequence of the scalar constraint $\hat{\mathcal{C}}$, see (12), and for certain types of matter fields also of the 3-diffeo constraint $\hat{\mathcal{C}}_i$ in (11), while the local Lorentz constraint $\hat{\mathcal{C}}_{ab}$ in (11) does not require entanglement. From this one can see that if the theory features anomalies due to the breaking of the 4D diffeomorphism symmetry, one cannot impose $\hat{\mathcal{C}}$ and $\hat{\mathcal{C}}_i$ as the Gupta–Bleuler-like conditions on the Hilbert space of the theory, and thus all subsequent results regarding the entanglement are void. In short, there cannot be any gauge protected entanglement if there is no relevant gauge symmetry to begin with. Nevertheless, if the theory features anomalies due to the breaking of the local Lorentz or any internal symmetries, while maintaining diffeomoprhism symmetry at the quantum level, the gauge protected entanglement will not be influenced by the anomaly.

5.4. Future lines of research

One of the main lines of future work would be to perform a detailed numerical analysis of $\mathrm{Tr}\,\hat{\rho}_M^2$ and the von Neumann entropy $S(\hat{\rho}_M)$ for the Hartle–Hawking state (either within the Regge, or some other QG model). The latter quantity, called the *entropy of entanglement*, represents the measure of the entanglement in *pure* and *bipartite* states [57], in our case between gravity and matter in the Hartle–Hawking state. The precise numerical deviation of the $\mathrm{Tr}\,\hat{\rho}_M^2$ from its maximal value 1 could indicate in which cases this new entanglement has relevant physical consequences. This way, it would be possible to determine the boundaries of validity of the assumption of the product gravity-matter states of the form $|G\rangle|M\rangle$, which has been up to now used in numerous studies (analogously to the case of determining the regimes in which two coherent states become effectively orthogonal). In connection to this, one could analyse in more detail quantitatively to what extent the gauge-protected gravity-matter entanglement

constrains the existence of macroscopic superpositions, and its effect to the quantum-to-classical transition (see the related work [3, 4, 9, 58–60]).

Further, studying the structure of the gauge-imposed entanglement for a tripartite system of gravity-matter-EM fields might bring qualitatively new effects. Unlike the case of pure bipartite states, where any two entangled states could be obtained from each other by local operations and classical communication (LOCC), thus forming a single class of entangled states and providing a unique measure of entanglement, the multipartite entanglement has a more complex structure. Indeed, in the tripartite case, in addition to the trivial classes of purely bipartite entanglement, say, $|a\rangle(|b_1c_1\rangle + |b_2c_2\rangle)$, genuine tripartite entanglement consists of a number of inequivalent classes of entangled states: in the simplest case of three qubits we have two classes of tripartite entanglement, represented by the states $|GHZ\rangle = (|0\,0\,0\rangle + |1\,1\,1\rangle)/\sqrt{2}$ and $|W\rangle = (|0\,0\,1\rangle + |0\,1\,0\rangle + |100\rangle)/\sqrt{3}$, which cannot be obtained from each other by the means of LOCC, but as soon as neither of the subsystems is a qubit, there exist infinitely many inequivalent classes [61].

It would also be interesting to see how other QG candidates incorporate the general gravity constraints regarding the entanglement with matter, in particular the string theory. Namely, perturbative string theory is formulated by manifestly breaking the gauge symmetry (a consequence of perturbative expansion of the gravitational field). The existence of the gravity-matter entanglement in, say Hartle–Hawking state, would then present a strong argument that the gauge symmetry can be restored in a tentative nonperturbative formulation of string theory. In connection to this, one could analyse the entanglement between different spacetime regions induced by the gauge-protected gravity-matter entanglement, and compare it to that present in theories based on the AdS/CFT correspondence and the holographic principle [7, 8]. Namely, entanglement is a property of a quantum state *with respect* to a particular factorisation of a composite system into its factor sub-systems. To illustrate this, consider a particle in a two-dimensional plane. Given orthogonal axes x and y of a 2D plane, the Hilbert space of the system is given by $\mathcal{H} = \mathcal{H}_x \otimes \mathcal{H}_y$, and the equal spatial superposition (for simplicity, we omit the overall normalisation constant) $|\varphi\rangle \sim (|a\rangle_x + |b\rangle_x)|0\rangle_y$, with $a, b \in \mathbb{R}$, is clearly separable, with respect to the given factorisation of $\mathcal{H}$. Nevertheless, with respect to *any other* factorisation of $\mathcal{H}$, defined by any other axes X and Y inducing the Hilbert-space factorisation $\mathcal{H} = \mathcal{H}_X \otimes \mathcal{H}_Y$, the system is entangled. As an example, for axes X and Y obtained by rotating x and y by $-\pi/4$, the same state of the system is maximally entangled, $|\varphi\rangle \sim (|a/\sqrt{2}\rangle_X|a/\sqrt{2}\rangle_Y + |b/\sqrt{2}\rangle_X|b/\sqrt{2}\rangle_Y)$ (for the entanglement in the second quantisation formalism, and its dependence on the choice of fundamental modes, see for example [62]). Following the above example, one might expect that the existence of the entanglement between gravity and matter would induce the entanglement between two generic spacetime regions (each containing a portion of both gravitational and matter degrees of freedom). Possible relationship between this, gauge-protected entanglement, and that present as a consequence of assumptions that do not explicitly rely on the existence of local Poincaré symmetry (holography and the AdS/CFT correspondence) would indicate interesting fundamental connections that could help breaching the long-standing gap between quantum mechanics and general relativity.

Finally, detecting gravity-matter entanglement in the experiment might not be that far from the reach of the current or the near-future technology, see [63] for a recent proposal of testing gravitational decoherence. Proposing, and possibly performing, experiments to distinguish different contributions of the gravitational interaction to the decoherence of matter, in particular the generic one based on the gauge symmetry constraints, presents a relevant direction of further research.

Acknowledgments

We would like to thank Rafael Sorkin for helpful suggestions and comments.

MV was supported by the project ON171031 of the Ministry of Education, Science and Technological Development of the Republic of Serbia, and the bilateral scientific cooperation between Portugal and Serbia through the project 'Quantum Gravity and Quantum Integrable Models—2015–2016', no. 451-03-01765/2014-09/24 supported by the Foundation for Science and Technology (FCT), Portugal, and the Ministry of Education, Science and Technological Development of the Republic of Serbia. NP acknowledges the support of SQIG—Security and Quantum Information Group, the IT project QbigD funded by FCT PEst-OE/EEI/LA0008/2013 and UID/EEA/50008/2013 and the bilateral scientific cooperation between Portugal and Serbia through the project 'Noise and measurement errors in multiparty quantum security protocols', no. 451-03-01765/2014-09/04 supported by the Foundation for Science and Technology (FCT), Portugal, and the Ministry of Education, Science and Technological Development of the Republic of Serbia.

The authors would also like to thank the Erwin Schrödinger International Institute for Mathematics and Physics (ESI Vienna), for the warm hospitality during the workshop 'Quantum Physics and Gravity' (29 May–30 June 2017) and partial support.

ORCID iDs

N Paunković https://orcid.org/0000-0002-9345-4321
M Vojinović https://orcid.org/0000-0001-6977-4870

References

[1] Karolyhazy F 1966 *Nuovo Cimento* A **42** 390
[2] Bassi A, Lochan K, Satin S, Singh T P and Ulbricht H 2013 *Rev. Mod. Phys.* **85** 471
[3] Pikovski I, Zych M, Costa F and Brukner Č 2011 *Nat. Commun.* **2** 505
[4] Pikovski I, Zych M, Costa F and Brukner Č 2015 *Nat. Phys.* **11** 668
[5] Oniga T and Wang C H T 2016 *Phys. Rev.* D **93** 044027
[6] Liu J, Sou C M and Wang Y 2016 *J. High Energy Phys.* JHEP10(2016)072
[7] Dong X 2016 *Nat. Commun.* **7** 12472
[8] Raamsdonk M V 2017 *Lectures on Gravity and Entanglement* (Oxford: World Scientific Publishing) pp 297–351
[9] Kay B S 1998 *Class. Quantum Grav.* **15** L89
[10] Kok P and Yurtsever U 2003 *Phys. Rev.* D **68** 085006
[11] Hamma A, Markopoulou F, Lloyd S, Caravelli F, Severini S and Markström K 2010 *Phys. Rev.* D **81** 104032
[12] Hamma A and Markopoulou F 2011 *New J. Phys.* **13** 095006
[13] Bassi A, Grossardt A and Ulbricht H 2017 *Class. Quantum Grav.* **34** 193002
[14] Hartle J B and Hawking S W 1983 *Phys. Rev.* D **28** 2960
[15] Blagojević M 2002 *Gravitation and Gauge Symmetries* (Bristol: Institute of Physics Publishing)
[16] Dirac P A M 1958 *Proc. R. Soc.* A **246** 333
[17] Dirac P A M 1964 *Lectures on Quantum Mechanics* (New York: Yeshiva University)
[18] Gupta S 1950 *Proc. Phys. Soc.* A **63** 681
[19] Bleuler K 1950 *Helv. Phys. Acta* **23** 567
[20] Nielsen M and Chuang I 2004 *Quantum Computation and Quantum Information* (*Cambridge Series on Information and the Natural Sciences*) (Cambridge: Cambridge University Press)
[21] Feldbrugge J, Lehners J L and Turok N 2018 *Phys. Rev.* D **97** 023509
[22] Rovelli C 2004 *Quantum Gravity* (Cambridge: Cambridge University Press)

[23] Rovelli C and Vidotto F 2014 *Covariant Loop Quantum Gravity: an Elementary Introduction to Quantum Gravity and Spinfoam Theory* (Cambridge: Cambridge University Press)
[24] Miković A and Vojinović M 2012 *Class. Quantum Grav.* **29** 165003
[25] Miković A 2013 *Rev. Math. Phys.* **25** 1343008
[26] Vojinović M 2016 *Phys. Rev.* D **94** 024058
[27] Hamber H W 2009 *Gen. Relativ. Gravit.* **41** 817
[28] Ambjørn J, Görlich A, Jurkiewicz J and Loll R 2012 *Phys. Rep.* **519** 127
[29] Freidel L and Livine E R 2006 *Phys. Rev. Lett.* **96** 221301
[30] Paunković N and Vojinović M 2016 *J. Phys.: Conf. Ser.* **701** 012035
[31] Regge T 1961 *Nuovo Cimento* **19** 558
[32] Miković A 2017 *Adv. Theor. Math. Phys.* **21** 631
[33] Miković A and Vojinović M 2015 *Europhys. Lett.* **110** 40008
[34] Miković A and Vojinović M 2014 Cosmological constant in a quantum gravity theory for a piecewise-linear spacetime (arXiv:1407.1124)
[35] Pipa F, Paunković N and Vojinović M 2018 Entanglement-induced deviation from the geodesic motion in quantum gravity (arXiv:1801.03207)
[36] Zurek W H 1981 *Phys. Rev.* D **24** 1516
[37] Schlosshauer M 2005 *Rev. Mod. Phys.* **76** 1267
[38] Penrose R 1996 *Gen. Relativ. Gravit.* **28** 581
[39] Everett H 1957 *Rev. Mod. Phys.* **29** 454
[40] DeWitt B S 2003 *The Global Approach to Quantum Field Theory* vol 1 (Oxford: Oxford University Press)
[41] Duff M J 1996 *Int. J. Mod. Phys.* A **11** 5623
[42] Lambert N 2012 *Ann. Rev. Nucl. Part. Sci.* **62** 285
[43] Becker K, Becker M and Schwarz J H 2007 *String Theory and M-Theory* (Cambridge: Cambridge University Press)
[44] DeWitt B S 1967 *Phys. Rev.* **160** 1113
[45] Arnowitt R, Deser S and Misner C W 2008 *Gen. Relativ. Gravit.* **40** 1997
[46] Feynman R, Morinigo F, Wagner W, Hatfield B and Pines D 2002 *Feynman Lectures on Gravitation* (Boulder, CO: Westview Press)
[47] Amelino-Camelia G 2012 Doubly-special relativity: facts, myths and some key open issues *Recent Developments in Theoretical Physics* ed S Ghosh and G Kar (Singapore: World Scientific)
[48] Hořava P 2009 *Phys. Rev.* D **79** 084008
[49] Niedermaier M and Reuter M 2006 *Living Rev. Relativ.* **9** 5
[50] Connes A 1994 *Noncommutative Geometry* (San Diego, CA: Academic)
[51] Sakellariadou M 2013 *J. Phys.: Conf. Ser.* **442** 012015
[52] Hehl F W, McCrea J D, Mielke E W and Ne'eman Y 1995 *Phys. Rep.* **258** 1
[53] Verlinde E 2011 *J. High Energy Phys.* JHEP04(2011)029
[54] Jacobson T 1995 *Phys. Rev. Lett.* **75** 1260
[55] Henson J 2009 The causal set approach to Quantum Gravity *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter* ed D Oriti (Cambridge: Cambridge University Press)
[56] Bombelli L, Lee J, Meyer D and Sorkin R D 1987 *Phys. Rev. Lett.* **59** 521
[57] Bennett C H, Bernstein H J, Popescu S and Schumacher B 1996 *Phys. Rev.* A **53** 2046
[58] Peres A and Rosen N 1964 *Phys. Rev.* **135** B1486
[59] Baym G and Ozawa T 2009 *Proc. Natl Acad. Sci.* **106** 3035
[60] Brukner Č 2017 On the quantum measurement problem *Quantum [Un]Speakables II: Half a Century of Bell's Theorem* ed R Bertlmann and A Zeilinger (Berlin: Springer) pp 95–117
[61] Dür W, Vidal G and Cirac J I 2000 *Phys. Rev.* A **62** 062314
[62] Vedral V 2003 *Open Phys.* **1** 289
[63] Pfister C, Kaniewski J, Tomamichel M, Mantri A, Schmucker R, McMahon N, Milburn G and Wehner S 2016 *Nat. Commun.* **7** 13022

Dynamical phase transitions at finite temperature from fidelity and interferometric Loschmidt echo induced metrics

Bruno Mera,[1,2,3] Chrysoula Vlachou,[2,4] Nikola Paunković,[2,4] Vítor R. Vieira,[1,2] and Oscar Viyuela[5,6]

[1]*CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

[2]*Departamento de Física, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

[3]*Instituto de Telecomunicações, 1049-001 Lisbon, Portugal*

[4]*Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

[5]*Department of Physics, Harvard University, Cambridge, Massachusetts 02318, USA*

[6]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

(Received 27 December 2017; revised manuscript received 6 March 2018; published 23 March 2018)

We study finite-temperature dynamical quantum phase transitions (DQPTs) by means of the fidelity and the interferometric Loschmidt echo (LE) induced metrics. We analyze the associated dynamical susceptibilities (Riemannian metrics), and derive analytic expressions for the case of two-band Hamiltonians. At zero temperature, the two quantities are identical, nevertheless, at finite temperatures they behave very differently. Using the fidelity LE, the zero-temperature DQPTs are gradually washed away with temperature, while the interferometric counterpart exhibits finite-temperature phase transitions. We analyze the physical differences between the two finite-temperature LE generalizations, and argue that, while the interferometric one is more sensitive and can therefore provide more information when applied to genuine quantum (microscopic) systems, when analyzing many-body macroscopic systems, the fidelity-based counterpart is a more suitable quantity to study. Finally, we apply the previous results to two representative models of topological insulators in one and two dimensions.

DOI: 10.1103/PhysRevB.97.094110

I. INTRODUCTION

Equilibrium phase transitions (PTs) are characterized by a nonanalytic behavior of relevant thermodynamic observables with respect to the change of temperature. Quantum phase transitions (QPTs) [1], traditionally described by Landau theory [2], occur when we adiabatically change a physical parameter of the system at zero temperature, i.e., the transition is driven by purely quantum fluctuations. A particularly interesting case arises when one studies QPTs in the context of topological phases of matter [3–7]. They are quite different from the standard QPTs since they do not involve symmetry breaking and are characterized by global order parameters [8]. These novel phases of matter, which include topological insulators and superconductors [9–11], have potentially many applications in emerging fields such as spintronics, photonics, or quantum computing. Although many of their remarkable properties have traditionally been studied at zero temperature, there has been a great effort to generalize these phases from pure to mixed states, and finite temperatures [12–24].

The real time evolution of closed quantum systems out of equilibrium has some surprising similarities with thermal phase transitions, as noticed by Heyl, Polkovnikov, and Kehrein [25]. They coined the term dynamical quantum phase transitions (DQPTs) to describe the nonanalytic behavior of certain dynamical observables after a sudden quench in one of the parameters of the Hamiltonian. Since then, the study of DQPTs became an active field of research, and a lot of progress has been achieved in comparing and connecting them to the equilibrium PTs [26–33]. Along those lines, there exist several studies of DQPTs for systems featuring nontrivial topological properties [34–39]. DQPTs have been experimentally observed in systems of trapped ions [40] and cold atoms in optical lattices [41]. The figure of merit in the study of DQPTs is the Loschmidt echo (LE) and its derivatives, which have been extensively used in the analysis of quantum criticality [42–46] and quantum quenches [47]. At finite temperature, generalizations of the zero-temperature LE were proposed, based on the mixed-state Uhlmann fidelity [46,48], and the interferometric mixed-state geometric phase [49,50]. For an alternative approach to finite-temperature DPTs, see [51,52]. Fidelity is a measure of state distinguishability, which has been employed numerous times in the study of PTs [43,53–56], while the interferometric mixed-state geometric phase was introduced in [57]. The two quantities are in general different and it comes as no surprise that they give different predictions for the finite-temperature behavior [50]: the fidelity LE does not show DPTs at finite temperatures, while the interferometric LE indicates their persistence at finite temperature. Thus, it remains unclear what the fate of DQPT at finite temperature truly is, and which of the two opposite predictions better captures the many-body nature of these PTs.

In this paper, we discuss the existence of finite-temperature dynamical phase transitions (DPTs) for the broad class of two-band Hamiltonians in terms of both the *fidelity* and the *interferometric LEs*. We derive analytic expressions for the metrics (susceptibilities) induced by the fidelity and the interferometric LEs, respectively, showing explicitly that the two approaches give different behaviors: the fidelity susceptibility shows a gradual disappearance of DPTs as the temperature increases, while the interferometric susceptibility indicates their persistence at finite temperature (consistent with

 ©2018 American Physical Society

recent studies on interferometric LE [49,50]). We analyze the reasons for such different behaviors. The fidelity LE quantifies state distinguishability in terms of measurements of physical properties, inducing a metric over the space of quantum states, while the interferometric LE quantifies the effects of quantum channels acting upon a state, inducing a pullback metric over the space of unitaries. Thus, we argue that the fidelity LE and its associated dynamical susceptibility are more suitable for the study of many-body systems, while the more sensitive interferometric counterparts are optimal when considering genuine microscopic quantum systems. In addition, interferometric experiments that are suitable for genuine (microscopic) quantum systems involve coherent superpositions of two states, which could be, in the case of many-body macroscopic systems, experimentally infeasible with current technology. To confirm our analysis for the fidelity LE, in the Appendix, we present quantitative results for the fidelity-induced first time derivative of the rate function in the case of the one-dimensional (1D) Su-Schrieffer-Heeger (SSH) topological insulator and the two-dimensional (2D) massive Dirac (MD) model of a Chern insulator [58].

II. DYNAMICAL (QUANTUM) PHASE TRANSITIONS AND ASSOCIATED SUSCEPTIBILITIES

The authors in [25] introduce the concept of DQPTs and illustrate their properties on the case of the transverse-field Ising model. They observe a similarity between the partition function of a quantum system in equilibrium, $Z(\beta) = \mathrm{Tr}(e^{-\beta H})$, and the overlap amplitude of some time-evolved initial quantum state $|\psi_i\rangle$ with itself, $G(t) = \langle\psi_i|e^{-iHt}|\psi_i\rangle$. During a temperature-driven PT, the abrupt change of the properties of the system is indicated by the nonanalyticity of the free energy density $f(\beta) = -\lim_{N\to\infty}\frac{1}{N}\ln Z(\beta)$ at the critical temperature (N being the number of degrees of freedom). It is then possible to establish an analogy with the case of the real-time evolution of a quantum system out of equilibrium, by considering the rate function

$$g(t) = -\frac{1}{N}\log|G(t)|^2, \tag{1}$$

where $|G(t)|^2$ is a mixed-state LE, as we detail below. The rate function $g(t)$ may exhibit nonanalyticities at some critical times t_c, after a quantum quench. This phenomenon is termed DPT.

We study DPTs for mixed states using the fidelity and the interferometric LEs. We first investigate the relation between the two approaches for DQPTs at zero temperature. More concretely, we perform analytical derivations of the corresponding susceptibilities in the general case of a family of static Hamiltonians, parametrized by some smooth manifold M, $\{H(\lambda) : \lambda \in M\}$.

A. DQPTs for pure states

At zero temperature, the LE $G(t)$ from (1) between the ground state for $\lambda = \lambda_i \in M$ and the evolved state with respect to the Hamiltonian for $\lambda = \lambda_f \in M$ is given by the fidelity between the two states

$$\mathcal{F}(t;\lambda_f,\lambda_i) \equiv |\langle\psi(\lambda_i)|e^{-itH(\lambda_f)}|\psi(\lambda_i)\rangle|. \tag{2}$$

For $\lambda_i = \lambda_f$, the fidelity is trivial since the system remains in the same state. Fixing $\lambda_i \equiv \lambda$ and $\lambda_f = \lambda + \delta\lambda$, with $\delta\lambda \ll 1$, in the $t \to \infty$ limit Eq. (2) is nothing but the familiar S matrix with an unperturbed Hamiltonian $H(\lambda)$ and an interaction Hamiltonian $V(\lambda)$, which is approximated by

$$V(\lambda) \equiv H(\lambda_f) - H(\lambda) \approx \frac{\partial H}{\partial\lambda^a}(\lambda)\delta\lambda^a. \tag{3}$$

After applying standard perturbation theory techniques (see Appendix), we obtain

$$\mathcal{F}(t;\lambda_f,\lambda) \approx 1 - \chi_{ab}(t;\lambda)\delta\lambda^a\delta\lambda^b, \tag{4}$$

where the dynamical susceptibility $\chi_{ab}(t;\lambda)$ is given by

$$\chi_{ab}(t;\lambda) = \int_0^t\int_0^t dt_2dt_1\Big(\frac{1}{2}\langle\{V_a(t_2),V_b(t_1)\}\rangle - \langle V_a(t_2)\rangle\langle V_b(t_1)\rangle\Big), \tag{5}$$

with $V_a(t,\lambda) = e^{itH(\lambda)}\partial H/\partial\lambda^a(\lambda)e^{-itH(\lambda)}$ and $\langle * \rangle = \langle\psi(\lambda)| * |\psi(\lambda)\rangle$. The family of symmetric tensors $\{ds^2(t) = \chi_{ab}(t,\lambda)d\lambda^a d\lambda^b\}_{t\in\mathbb{R}}$ defines a family of metrics in the manifold M, which can be seen as pullback metrics of the Bures metric (Fubini-Study metric) in the manifold of pure states [59]. Specifically, at time t, the pullback is given by the map $\Phi_t : \lambda_f \mapsto e^{-itH(\lambda_f)}|\psi(\lambda)\rangle\langle\psi(\lambda)|e^{itH(\lambda_f)}$, evaluated at $\lambda_f = \lambda$.

B. Generalizations at finite temperatures

The generalization of DQPTs to mixed states is not unique. There are several ways to construct a LE for a general density matrix. In what follows, we derive two finite-temperature generalizations, such that they have the same zero-temperature limit.

1. Fidelity Loschmidt echo at $T > 0$

First, we introduce the *fidelity LE* between the state $\rho(\beta;\lambda_i) = e^{-\beta H(\lambda_i)}/\mathrm{Tr}\{e^{-\beta H(\lambda_i)}\}$ and the one evolved by the unitary operator $e^{-itH(\lambda_f)}$ as

$$\mathcal{F}(t,\beta;\lambda_i,\lambda_f) = F\big(\rho(\beta;\lambda_i), e^{-itH(\lambda_f)}\rho(\beta;\lambda_i)e^{itH(\lambda_f)}\big), \tag{6}$$

where $F(\rho,\sigma) = \mathrm{Tr}\{\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\}$ is the quantum fidelity between arbitrary mixed states ρ and σ. For λ_f close to $\lambda_i = \lambda$, we can write

$$\mathcal{F}(t,\beta;\lambda_f,\lambda) \approx 1 - \chi_{ab}(t,\beta;\lambda)\delta\lambda^a\delta\lambda^b, \tag{7}$$

with $\chi_{ab}(t,\beta,\lambda)$ being the *dynamical fidelity susceptibility* (DFS). Notice that $\lim_{\beta\to\infty}\chi_{ab}(t,\beta;\lambda) = \chi_{ab}(t;\lambda)$, where $\chi_{ab}(t;\lambda)$ is given by Eq. (5). At time t and inverse temperature β, we have a map $\Phi_{(t,\beta)} : \lambda_f \mapsto e^{-itH(\lambda_f)}\rho(\beta;\lambda)e^{itH(\lambda_f)}$. The two-parameter family of metrics defined by $ds^2(\beta,t) = \chi_{ab}(t,\beta;\lambda)d\lambda^a d\lambda^b$ is the pullback by $\Phi_{(t,\beta)}$ of the Bures metric on the manifold of full-rank density operators, evaluated at $\lambda_f = \lambda$ (see Appendix).

The fidelity LE is closely related to the Uhlmann connection: $F(\rho_1,\rho_2)$ equals the overlap between purifications W_1 and W_2, $\langle W_1,W_2\rangle = \mathrm{Tr}\{W_1^\dagger W_2\}$, satisfying discrete parallel transport condition (see, for instance, [60]).

2. *Interferometric Loschmidt echo at $T > 0$*

Here, we consider an alternative definition of the LE for mixed states [$G(t)$ from Eq. (1)]. In particular, we define the *interferometric LE* as

$$\mathcal{L}(t,\beta;\lambda_f,\lambda_i) = \left|\frac{\mathrm{Tr}\{e^{-\beta H(\lambda_i)}e^{itH(\lambda_i)}e^{-itH(\lambda_f)}\}}{\mathrm{Tr}\{e^{-\beta H(\lambda_i)}\}}\right|. \tag{8}$$

The $e^{itH(\lambda_i)}$ factor does not appear at zero temperature since it just gives a phase which is canceled by taking the absolute value. This differs from previous treatments in the literature [36] (see Sec. 5.5.4 of [59], where the variation of the interferometric phase, $\mathrm{Tr}\{\rho_0 e^{-itH}\}$, exposes this structure). However, it is convenient to introduce it in order to have the usual form of the perturbation expansion, as will become clear later.

For λ_f close to $\lambda_i = \lambda$, we get

$$\mathcal{L}(t,\beta;\lambda_f,\lambda) \approx \left|\frac{\mathrm{Tr}\{e^{-\beta H(\lambda)}\mathcal{T}e^{-i\int_0^t dt' V_a(t,\lambda)\delta\lambda^a}\}}{\mathrm{Tr}\{e^{-\beta H(\lambda)}\}}\right|, \tag{9}$$

so that the perturbation expansion goes as in Eq. (5), yielding

$$\mathcal{L}(t,\beta;\lambda_f,\lambda) \approx 1 - \tilde{\chi}_{ab}(t,\beta;\lambda)\delta\lambda^a\delta\lambda^b, \tag{10}$$

with the dynamical susceptibility given by

$$\tilde{\chi}_{ab}(t,\beta;\lambda) = \int_0^t\int_0^t dt_2 dt_1 \left(\frac{1}{2}\langle\{V_a(t_2),V_b(t_1)\}\rangle - \langle V_a(t_2)\rangle\langle V_b(t_1)\rangle\right), \tag{11}$$

where $\langle * \rangle = \mathrm{Tr}\{e^{-\beta H(\lambda)}*\}/\mathrm{Tr}\{e^{-\beta H(\lambda)}\}$. Notice that Eqs. (11) and (5) are formally the same with the average over the ground state replaced by the thermal average. This justifies the extra $e^{itH(\lambda_i)}$ factor. Since this susceptibility comes from the interferometric LE, we call it *dynamical interferometric susceptibility* (DIS). The quantity $ds^2(\beta,t) = \tilde{\chi}_{ab}(t,\beta;\lambda)d\lambda^a d\lambda^b$ defines a two-parameter family of metrics over the manifold M, except that they cannot be seen as pullbacks of metrics on the manifold of density operators with full rank. However, it can be interpreted as the pullback by a map from M to the unitary group associated with the Hilbert space of a particular Riemannian metric. For a detailed analysis, see the Appendix. Additionally, we point out that this version of LE is related to the interferometric geometric phase introduced by Sjöqvist *et al.* [57,61].

C. Two-band systems

Many representative examples of topological insulators and superconductors can be described by effective two-band Hamiltonians. Therefore, we derive closed expressions of the previously introduced dynamical susceptibilities for topological systems within this class.

The general form of such Hamiltonians is $\{H(\lambda) = \vec{x}(\lambda)\cdot\vec{\sigma} : \lambda \in M\}$, where $\vec{\sigma}$ is the Pauli vector. The interaction Hamiltonian $V(\lambda)$, introduced in Eq. (2), casts the form

$$V(\lambda) \approx \left(\frac{\partial\vec{x}}{\partial\lambda^a}\cdot\vec{\sigma}\right)\delta\lambda^a.$$

It is convenient to decompose $\partial\vec{x}/\partial\lambda^a$ into one component perpendicular to $\vec{x}$ and one parallel to it:

$$\frac{\partial\vec{x}}{\partial\lambda^a} = \left(\frac{\partial\vec{x}}{\partial\lambda^a}\right)^{\perp} + \left(\frac{\partial\vec{x}}{\partial\lambda^a}\right)^{\parallel} = \vec{t}_a + \vec{n}_a.$$

The first term is tangent, in $\mathbb{R}^3$, at $\vec{x}(\lambda)$, to a sphere of constant radius $r = |\vec{x}(\lambda)|$. Hence, this kind of perturbation does not change the spectrum of H, only its eigenbasis. The second term is a variation of the length of $\vec{x}$ and, hence, it changes the spectrum of H, while keeping the eigenbasis fixed. The DFS and the DIS are given by (for the details of the derivation, see Appendix)

$$\chi_{ab} = \tanh^2(\beta|\vec{x}(\lambda)|)\frac{\sin^2(|\vec{x}(\lambda)|t)}{|\vec{x}(\lambda)|^2}\vec{t}_a\cdot\vec{t}_b \tag{12}$$

$$\tilde{\chi}_{ab} = \frac{\sin^2(|\vec{x}(\lambda)|t)}{|\vec{x}(\lambda)|^2}\vec{t}_a\cdot\vec{t}_b + t^2[1-\tanh^2(\beta|\vec{x}(\lambda)|)]\vec{n}_a\cdot\vec{n}_b. \tag{13}$$

While the DIS (13) depends on the variation of both the spectrum and the eigenbasis of the Hamiltonian, the DFS (12) depends only on the variations which preserve the spectrum, i.e., changes in the eigenbasis. This is very remarkable; in general, the fidelity between two quantum states, being their distinguishability measure, does depend on both the variations of the spectrum and the eigenbasis. In our particular case of a quenched system, the eigenvalues are preserved [see Eq. (6)], as the system is subject to a unitary evolution. The tangential components of both susceptibilities are modulated by the function $\sin^2(Et)/E^2$, where E is the gap. This captures the *Fisher zeros*, i.e., the zeros of the (dynamical) partition function which here is given by the fidelity $\mathcal{F}$ from (6) (see [62–64]). Observe that whenever $t = (2n+1)\pi/2E$, $n \in \mathbb{Z}$, this factor is maximal and, hence, both LEs decrease abruptly. The difference between the two susceptibilities is given by

$$\tilde{\chi}_{ab} - \chi_{ab} = [1-\tanh^2(\beta|\vec{x}(\lambda)|)] \times\left(\frac{\sin^2(|\vec{x}(\lambda)|t)}{|\vec{x}(\lambda)|^2}\vec{t}_a\cdot\vec{t}_b + t^2\vec{n}_a\cdot\vec{n}_b\right).$$

The quantity $[1-\tanh^2(\beta E)]$ is nothing but the static susceptibility (see [65]). Therefore, the difference between DIS and DFS is modulated by the static susceptibility at finite temperature.

To illustrate the relationship between the two susceptibilities, in Fig. 1 we plotted the modulating function for the tangential components of both. We observe that at zero temperature they coincide. As the temperature increases, in the case of the fidelity LE, the gap-vanishing points become less important. On the contrary, for the interferometric LE, the associated tangential part of the susceptibility does not depend on temperature, thus, the gap-vanishing points remain prominent. The DFS from Eq. (12) thus predicts gradual smearing of critical behavior, consistent with previous findings that showed the absence of phase transitions at finite temperatures in the static case [14,15]. The DIS from Eq. (13) has a tangential term that is not coupled to the temperature, persisting at higher temperatures and giving rise to abrupt changes in the finite-temperature system's behavior. This is also consistent with

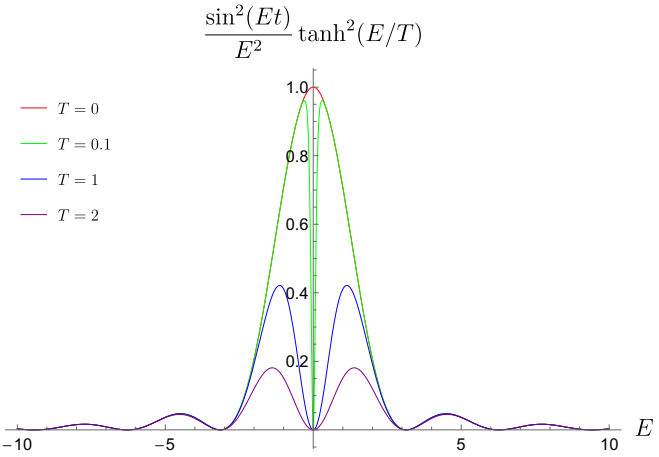


FIG. 1. The susceptibility modulating function for the tangential components at $t = 1$.

previous studies in the literature, where DPTs were found even at finite temperatures [49,50]. Additionally, the interferometric LE depends on the normal components of the variation of $\vec{x}$. In other words, the finite-temperature phase transitions inferred by the behavior of the interferometric LE occur due to the change of the parameters of the Hamiltonian and not due to temperature.

D. Comparing the two approaches

The above analysis of the two dynamical susceptibilities (metrics) reflects the essential difference between the two distinguishability measures, one based on the fidelity, the other on interferometric experiments. From the quantum information theoretical point of view, the two quantities can be interpreted as distances between *states*, or between *processes*, respectively. The Hamiltonian evaluated at a certain point of parameter space M defines the macroscopic phase. Associated to it we have thermal states and unitary processes. The fidelity LE is obtained from the Bures distance between a thermal state ρ_1 in phase 1 and the one obtained by unitarily evolving this state, $\rho_2 = U_2\rho_1 U_2^\dagger$, with U_2 associated to phase 2. Given a thermal state ρ_1 prepared in phase 1, the interferometric LE is obtained from the distance between two unitary processes U_1 and U_2 (defined modulo a phase factor), associated to phases 1 and 2.

The quantum fidelity between two states is in fact the classical fidelity between the probability distributions obtained by performing an optimal measurement on them. Measuring an observable M on the two states ρ_1 and ρ_2, one obtains the probability distributions $\{p_1(i)\}$ and $\{p_2(i)\}$, respectively. The quantum fidelity F_Q between the two states ρ_1 and ρ_2 is bounded by the classical fidelity F_C between the probability distributions $\{p_1(i)\}$ and $\{p_2(i)\}$, $F_Q(\rho_1,\rho_2) = \mathrm{Tr}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}} \leqslant \sum_i \sqrt{p_1(i)p_2(i)} = F_c(p_1(i),p_2(i))$, such that the equality is obtained by measuring an *optimal observable*, given by $M_{\mathrm{op}} = \rho_1^{-1/2}\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\rho_1^{-1/2}$ (note that optimal observable is not unique). For that reason, one can argue that the fidelity is capturing all order parameters (i.e., measurements) through its optimal observables M_{op}. Fidelity-induced distances, the *Bures distance* $D_B(\rho_1,\rho_2) = \sqrt{2[1 - F_Q(\rho_1,\rho_2)]}$, the *sine distance* $D_S(\rho_1,\rho_2) = \sqrt{1 - F_Q^2(\rho_1,\rho_2)}$ and the *F distance* $D_F(\rho_1,\rho_2) = 1 - F_Q(\rho_1,\rho_2)$ satisfy the following set of inequalities:

$$D_F(\rho_1,\rho_2) \leqslant D_T(\rho_1,\rho_2) \leqslant D_S(\rho_1,\rho_2) \leqslant D_B(\rho_1,\rho_2),$$

where the *trace distance* is given by $D_T(\rho_1,\rho_2) = \frac{1}{2}\mathrm{Tr}|\rho_1 - \rho_2|$. In other words, the fidelity-induced distances and the trace distance establish the same order on the space of quantum states. This is important, as the trace distance is giving the optimal value for the success probability in ambiguously discriminating in a *single-shot measurement* between two *a priori* equally probable states ρ_1 and ρ_2, given by the so-called Helstrom bound $P_H(\rho_1,\rho_2) = [1 + D_T(\rho_1,\rho_2)]/2$ [66].

On the other hand, the interferometric phase is based on some interferometric experiment to distinguish two states $\rho_1 = \sum_i r_i|i\rangle\langle i|$ and $\rho_2 = U_2\rho_1 U_2^\dagger$: it measures how the intensities at the outputs of the interferometer are affected by applying U_2 to only one of its arms [57]. Therefore, to set up such an experiment, one does not need to know the state ρ_1 that enters the interferometer, as only the knowledge of U_2 is required. Note that this does not mean that the output intensities do not depend on the interferometric LE: indeed, the inner product $\langle U_1, U_2\rangle_{\rho_1}$ is defined with respect to the state ρ_1. This is a different type of experiment, not based on the observation of any physical property of a system. It is analogous to comparing two masses with weighing scales, which would show the same difference of $\Delta m = m_1 - m_2$, regardless of how large the two masses m_1 and m_2 are. For that reason, interferometric distinguishability is more sensitive than the fidelity (fidelity depends on more information, not only how much the two states are different, but in what aspects this difference is observable). Indeed, the interferometric LE between ρ_1 and ρ_2 can be written as the overlap $L(\rho_1,\rho_2) = |\langle\rho_1|\rho_2\rangle|$ between the purifications $|\rho_1\rangle = \sum_i \sqrt{r_i}|i\rangle|i\rangle$ and $|\rho_2\rangle = (U \otimes I)|\rho_1\rangle$. On the other hand, the fidelity satisfies $F(\rho_1,\rho_2) = \max_{|\psi\rangle,|\varphi\rangle} |\langle\psi|\varphi\rangle|$, where $|\psi\rangle$ and $|\varphi\rangle$ are purifications of ρ_1 and ρ_2, respectively, i.e., $L(\rho_1,\rho_2) \leqslant F(\rho_1,\rho_2)$. Moreover, what one does observe in interferometric experiments are the mentioned output intensities, i.e., one needs a number of identical systems prepared in the same state to obtain results in interferometric measurements. This additionally explains why interferometric LE is more sensible than the fidelity one, as the latter is based on the observations performed on single systems. The fact that interferometric LE is more sensitive than the fidelity LE is consistent with the result that the former is able to capture the changes of some of the system's features at finite temperatures (thus they predict DPTs), while the latter cannot.

In terms of experimental feasibility, the fidelity is more suitable for the study of many-body macroscopic systems and phenomena, while the interferometric measurements provide a more detailed information on genuinely quantum (microscopic) systems. Finally, interferometric experiments involve coherent superpositions of two states. Therefore, when applied to many-body systems, one would need to create genuine Schrödinger catlike states, which goes beyond the current, and any foreseeable, technology (and could possibly be forbidden by more fundamental laws of physics; see for example objective collapse theories [67]).

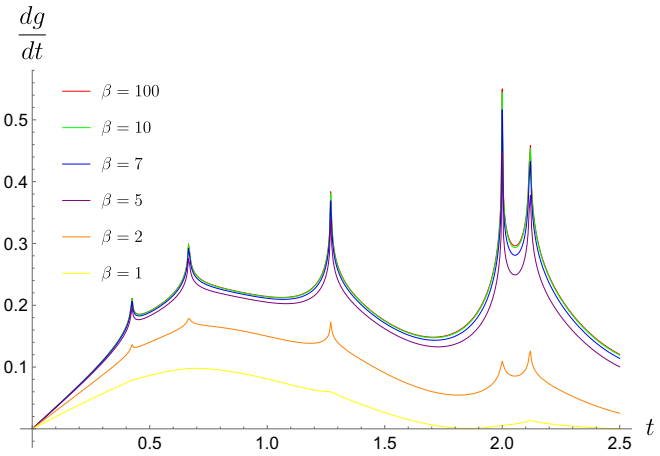


FIG. 2. We plot the time derivative of the rate function dg/dt as a function of time for different values of the inverse temperature $\beta = 1/T$. We consider a quantum quench from a trivial phase ($m = 1.2$) to a topological phase ($m = 0.8$).

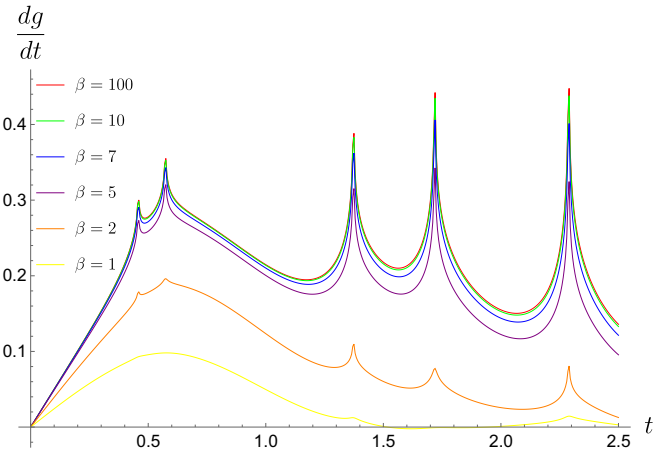


FIG. 3. The time derivative of the rate function dg/dt as a function of time for different values of the inverse temperature. The quench is from a topological ($m = 0.8$) to a trivial phase ($m = 1.2$).

III. DPTs OF TOPOLOGICAL INSULATORS AT FINITE TEMPERATURES

Our general study of two-band Hamiltonians showed that the fidelity-induced LE predicts a gradual smearing of DPTs with temperature. In order to test this result, we study the fidelity LE on concrete examples of two topological insulators (as noted in the main text, the analogous study for the interferometric LE on concrete examples has already been performed, and is consistent with our findings [49,50]). In particular, we present quantitative results obtained for the first derivative of the rate function dg/dt, where $g(t) = -\frac{1}{N}\log\mathcal{F}$, and

$$\mathcal{F}(t,\beta;\lambda_i,\lambda_f) = F\left(\rho(\beta;\lambda_i), e^{-itH(\lambda_f)}\rho(\beta;\lambda_i)e^{itH(\lambda_f)}\right).$$

The fidelity F is obtained by taking the product of the single-mode fidelities, each of which has the form

$$F\left(\rho(\beta;\lambda_i), e^{-itH(\lambda_f)}\rho(\beta;\lambda_i)e^{itH(\lambda_f)}\right) = \sqrt{\frac{1+\cosh^2(\beta E_i)+\sinh^2(\beta E_i)\{\cos(2E_f t)+[1-\cos(2E_f t)](\vec{n}_i\cdot\vec{n}_f)^2\}}{2\cosh^2(\beta E_i)}},$$

with $H_a = E_a\vec{n}_a\cdot\vec{\sigma}$ and $a = i,f$. This expression can be obtained by using Eq. (A4) and the result found in the Supplemental Material of [14]. The quantity dg/dt is the figure of merit in the study of the DQPTs, therefore, we present the respective results that confirm the previous study: the generalization of the LE with respect to the fidelity shows the absence of finite-temperature dynamical PTs. We consider two paradigmatic models of topological insulators, namely, the SSH [68] and the MD [58] models.

A. SSH model (1D)

The SSH model was introduced in [68] to describe polyacetylene, and it was later found to describe diatomic polymers [69]. In momentum space, the Hamiltonian for this model is of the form $H(k,m) = \vec{x}(k,m)\cdot\vec{\sigma}$, with m being the parameter that drives the static PT. The vector $\vec{x}(k,m)$ is given by

$$\vec{x}(k,m) = (m+\cos(k), \sin(k), 0).$$

By varying m we find two distinct topological regimes. For $m < m_c = 1$ the system is in a nontrivial phase with winding number 1, while for $m > m_c = 1$ the system is in a topologically trivial phase with winding number 0.

We consider both cases in which we go from a trivial to a topological phase and vice versa (Figs. 2 and 3, respectively). We notice that nonanalyticities of the first derivative appear

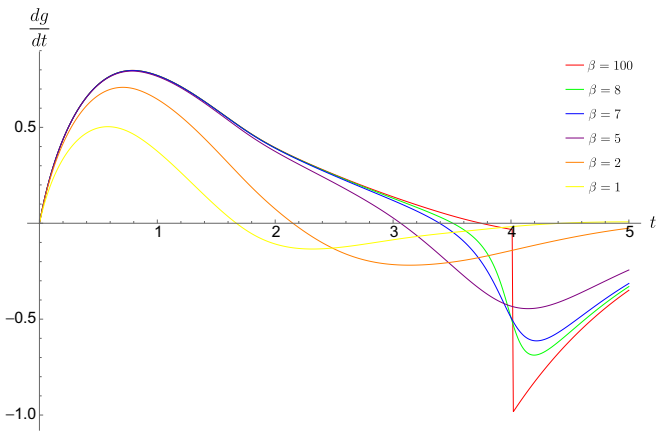


FIG. 4. The time derivative of the rate function dg/dt as a function of time for different values of the inverse temperature. We quench the system from a trivial to a topological regime (regimes from I to II and from IV to III).

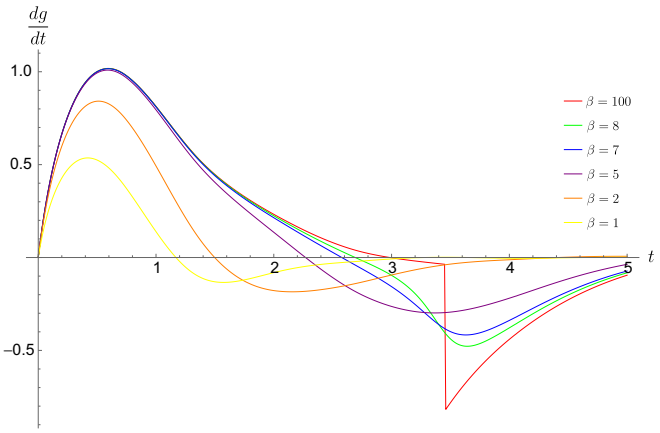


FIG. 5. The time derivative of the rate function dg/dt as a function of time for different values of the inverse temperature. The quench is from a topological to a trivial regime (regimes from II to I and from III to IV).

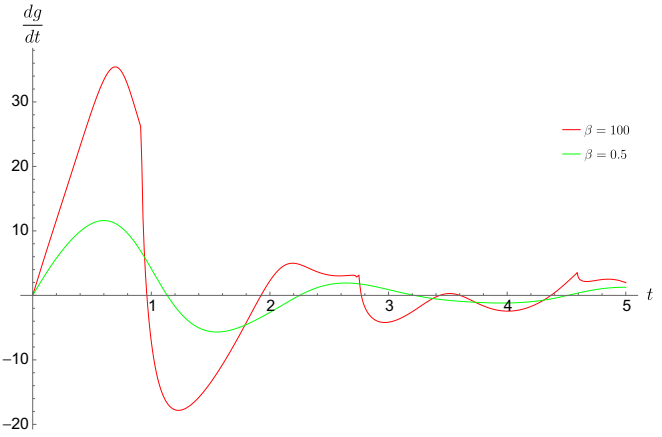


FIG. 7. The time derivative of the rate function dg/dt as a function of time for different values of the inverse temperature, in the case that we quench the system from a trivial to a topological regime (regimes from I to III and from IV to II).

at zero temperature, and they are smeared out for higher temperatures.

B. MD model (2D)

The massive Dirac model (MDM) captures the physics of a 2D Chern insulator [58], and shows different topologically distinct phases. In momentum space, the Hamiltonian for the MDM is of the form $H(\vec{k},m) = \vec{x}(\vec{k},m) \cdot \vec{\sigma}$, with m being the parameter that drives the static PT. The vector $\vec{x}(\vec{k},m)$ is given by

$$\vec{x}(\vec{k},m) = (\sin(k_x), \sin(k_y), m - \cos(k_x) - \cos(k_y)).$$

By varying m, we find four different topological regimes:

(i) For $-\infty < m < m_{c_1} = -2$ it is trivial (the Chern number is zero): regime I.

(ii) For $-2 = m_{c_1} < m < m_{c_2} = 0$ it is topological (the Chern number is -1): regime II.

(iii) For $0 = m_{c_2} < m < m_{c_3} = 2$ it is topological (the Chern number is $+1$): regime III.

Iiv) For $2 = m_{c_3} < m < \infty$ it is trivial (the Chern number is zero): regime IV.

In Figs, 4, 5, and 6 we plot the first derivative of the rate function $g(t)$, as a function of time for different temperatures. We only consider quenches that traverse a single-phase transition point.

We observe that at zero temperature there exist nonanalytycities at the critical times: the signatures of DQPTs. As we increase the temperature, these nonanalyticities are gradually smeared out, resulting in smooth curves for higher finite temperatures. We note that the peak of the derivative $dg(t)/dt$ is drifted when increasing the temperature, in analogy to the usual drift of nondynamical quantum phase transitions at finite temperature [70].

Next, we proceed by considering the cases in which we cross two phase transition points, as shown in Figs. 7 and 8.

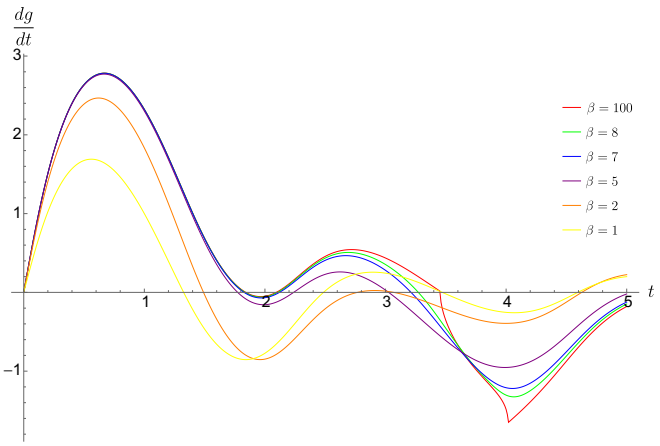


FIG. 6. The time derivative of the rate function dg/dt as a function of time for different inverse temperatures. The quantum quench is from a topological to a topological regime (regimes from II to III and vice versa).

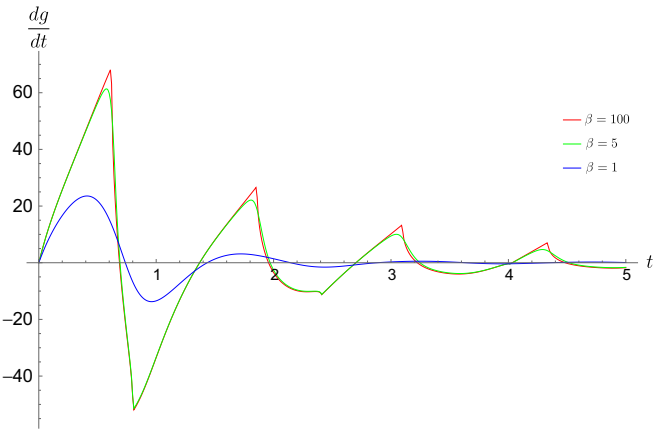


FIG. 8. The time derivative of the rate function dg/dt as a function of time for different values of the inverse temperature. The system is quenched from a topological to a trivial regime (regimes from III to I and from II to IV).

At zero temperature, we obtain a nonanalytic behavior, which gradually disappears for higher temperatures.

Finally, we have also studied the case in which we move inside the same topological regime from left to right and vice versa. We obtained smooth curves without nonanalyticities, which we omit for the sake of briefness.

IV. CONCLUSIONS

We analyzed the fidelity and the interferometric generalizations of the LE for general mixed states, and applied them to the study of finite-temperature DPTs in topological systems. We showed that the dynamical fidelity susceptibility is the pullback of the Bures metric in the *space of density matrices*, i.e., states. On the other hand, the dynamical interferometric susceptibility is the pullback of a metric in the *space of unitaries* (i.e., quantum channels).

The difference between the two metrics reflects the fact that the fidelity is a measure of the state distinguishability between two *given* states ρ and σ in terms of observations, while the "interferometric distinguishability" quantifies how a quantum channel (a unitary U) changes an *arbitrary* state ρ to $U\rho U^\dagger$.

Therefore, while the "interferometric distinguishability" is in general more sensitive, and thus appropriate for the study of genuine (microscopic) systems, it is the fidelity that is the most suitable for the study of many-body system phases.

Moreover, interferometric experiments involve coherent superpositions of two states, which for many-body systems would require creating and manipulating genuine Schrödinger catlike states. This seems to be experimentally beyond current technology.

We presented analytic expressions for the dynamical susceptibilities in the case of two-band Hamiltonians. At finite temperature, the fidelity LE indicates gradual disappearance of the zero-temperature DQPTs, while the interferometric LE predicts finite-temperature DPTs. We have performed finite-temperature study on two representatives of topological insulators: the 1D SSH and the 2D MD models. In perfect agreement with the general result, the fidelity-induced first derivatives gradually smear down with temperature, not exhibiting any critical behavior at finite temperatures. This is consistent with recent studies of 1D symmetry-protected topological phases at finite temperatures [14,15]. On the contrary, the interferometric LE exhibits critical behavior even at finite temperatures (confirming previous studies on DPTs [49,50]).

Note added. Recently, we became aware of a recent related work [39].

ACKNOWLEDGMENTS

B.M. and C.V. acknowledge the support from DP-PMI and FCT (Portugal) through the Grants No. SFRH/BD/52244/2013 and No. PD/BD/52652/2014, respectively. N.P. acknowledges the IT project QbigD funded by FCT Projects No. PEst-OE/EEI/LA0008/2013, No. UID/EEA/50008/2013, and the Confident Project No. PTDC/EEI-CTP/4503/2014. Support from FCT (Portugal) through Grant No. UID/CTM/04540/2013 is also acknowledged. O.V. thanks Fundación Rafael del Pino, Fundación Ramón Areces, and RCC Harvard.

APPENDIX: ANALYTICAL DERIVATION OF THE DYNAMICAL SUSCEPTIBILITIES

1. Zero-temperature case

Let $\mathscr{H}$ be a Hilbert space. Suppose we have a family of Hamiltonians $\{H(\lambda) : \lambda \in M\}$ where M is a smooth compact manifold of the Hamiltonian's parameters. We assume that aside from a closed finite subset of M, $C = \{\lambda_i\}_{i=1}^n \subset M$, the Hamiltonian is gapped and the ground-state subspace is one dimensional. Locally, on $M - C$, we can find a ground state (with unit norm) described by $|\psi(\lambda)\rangle$. Take $\lambda_i \in C$, and let U be an open neighborhood containing λ_i. Of course, for sufficiently small U, on the open set $U - \{\lambda_i\}$ one can find a smooth assignment $\lambda \mapsto |\psi(\lambda)\rangle$. Consider a curve $[0,1] \ni s \mapsto \lambda(s) \in U$, with initial condition $\lambda(0) = \lambda_0$, such that $\lambda(s_0) = \lambda_i$ for some $s_0 \in [0,1]$. The family of Hamiltonians $H(s) := H(\lambda(s))$ is well defined for every $s \in [0,1]$. The family of states $|\psi(s)\rangle \equiv |\psi(\lambda(s))\rangle$ is well defined for $s \neq s_0$ and so is the ground-state energy

$$E(s) := \langle\psi(s)|H(s)|\psi(s)\rangle.$$

The overlap

$$\mathcal{A}(s) := \langle\psi(0)|\exp[-itH(s)]|\psi(0)\rangle$$

is well defined. We can write

$$\exp[-itH(s)] = \exp[-itH(0)]T\exp\left\{-i\int_0^t d\tau\, V(s,\tau)\right\}.$$

If we take a derivative with respect to t of the equation, we find

$$H(s) = H(0) + \exp[-itH(0)]V(s,t)\exp[itH(0)]$$

so

$$V(s,t) = \exp[itH(0)][H(s) - H(0)]\exp[-itH(0)].$$

We can now write, since $|\psi(0)\rangle$ is an eigenvector of $H(0)$,

$$\mathcal{A}(s) = e^{-itE(0)}\langle\psi(0)|T\exp\left\{-i\int_0^t d\tau\, V(s,\tau)\right\}|\psi(0)\rangle.$$

We now perform an expansion of the overlap

$$\langle\psi(0)|T\exp\left\{-i\int_0^t d\tau\, V(s,\tau)\right\}|\psi(0)\rangle$$

in powers of s. Notice that

$$\begin{aligned} &T\exp\left\{-i\int_0^t d\tau\, V(s,\tau)\right\} \\ &\quad = I - i\int_0^t d\tau\, V(s,\tau) \\ &\qquad - \frac{1}{2}\int_0^t\int_0^t d\tau_2 d\tau_1 T\{V(s,\tau_2)V(s,\tau_1)\} + \cdots \end{aligned}$$

and hence

$$\frac{d}{ds}\left(T\exp\left\{-i\int_0^t d\tau\, V(s,\tau)\right\}\right)\Big|_{s=0} = -i\int_0^t d\tau\frac{\partial V}{\partial s}(0,\tau)$$

and

$$\frac{d^2}{ds^2}\left(T\exp\left\{-i\int_0^t d\tau\, V(s,\tau)\right\}\right)\Big|_{s=0} = -i\int_0^t d\tau\frac{\partial^2 V}{\partial s^2}(0,\tau) - \int_0^t\int_0^t d\tau_2 d\tau_1 T\left\{\frac{\partial V}{\partial s}(0,\tau_2),\frac{\partial V}{\partial s}(0,\tau_1)\right\}.$$

Therefore,

$$\langle\psi(0)|T\exp\left\{-i\int_0^t d\tau\, V(s,\tau)\right\}|\psi(0)\rangle = 1 - is\langle\psi(0)|\int_0^t d\tau\frac{\partial V}{\partial s}(0,\tau)|\psi(0)\rangle + \frac{s^2}{2}\Big[-i\langle\psi(0)|\int_0^t d\tau\frac{\partial^2 V}{\partial s^2}(0,\tau) \times|\psi(0)\rangle - \langle\psi(0)|\int_0^t\int_0^t d\tau_2 d\tau_1 T\left\{\frac{\partial V}{\partial s}(0,\tau_2),\frac{\partial V}{\partial s}(0,\tau_1)\right\}|\psi(0)\rangle\Big] + O(s^3).$$

Thus, by using the identity $\theta(\tau)+\theta(-\tau)=1$ of the Heaviside theta function, we obtain

$$|\mathcal{A}(s)|^2 = 1 - s^2\Big[\int_0^t\int_0^t d\tau_2 d\tau_1\langle\psi(0)|\frac{1}{2}\left\{\frac{\partial V}{\partial s}(0,\tau_2),\frac{\partial V}{\partial s}(0,\tau_1)\right\}|\psi(0)\rangle - \langle\psi(0)|\frac{\partial V}{\partial s}(0,\tau_2)|\psi(0)\rangle\langle\psi(0)|\frac{\partial V}{\partial s}(0,\tau_1)|\psi(0)\rangle\Big] + O(s^3).$$

If we denote the expectation value $\langle\psi(0)|*|\psi(0)\rangle \equiv \langle *\rangle$, we can write

$$|\mathcal{A}(s)|^2 = 1 - s^2\int_0^t\int_0^t d\tau_2 d\tau_1\left[\left\langle\frac{1}{2}\left\{\frac{\partial V}{\partial s}(0,\tau_2),\frac{\partial V}{\partial s}(0,\tau_1)\right\}\right\rangle - \left\langle\frac{\partial V}{\partial s}(0,\tau_2)\right\rangle\left\langle\frac{\partial V}{\partial s}(0,\tau_1)\right\rangle\right] + O(s^3) = 1 - \chi s^2 + O(s^3),$$

where

$$\chi \equiv \int_0^t\int_0^t d\tau_2 d\tau_1\left[\left\langle\frac{1}{2}\left\{\frac{\partial V}{\partial s}(0,\tau_2),\frac{\partial V}{\partial s}(0,\tau_1)\right\}\right\rangle - \left\langle\frac{\partial V}{\partial s}(0,\tau_2)\right\rangle\left\langle\frac{\partial V}{\partial s}(0,\tau_1)\right\rangle\right]$$

is the dynamical susceptibility and is naturally non-negative. In fact, defining $V_a(\tau) = e^{i\tau H(0)}\partial H/\partial\lambda^a(\lambda_0)e^{-i\tau H(0)}$ such that, by the chain rule

$$\frac{\partial V}{\partial s}(0,\tau) = V_a(\tau)\frac{\partial\lambda^a}{\partial s}(0),$$

we can write

$$\chi = g_{ab}(\lambda_0)\frac{\partial\lambda^a}{\partial s}(0)\frac{\partial\lambda^b}{\partial s}(0),$$

with the metric tensor given by

$$g_{ab}(\lambda_0) = \int_0^t\int_0^t d\tau_2 d\tau_1\left[\left\langle\frac{1}{2}\{V_a(\tau_2),V_b(\tau_1)\}\right\rangle - \langle V_a(\tau_2)\rangle\langle V_b(\tau_1)\rangle\right]. \tag{A1}$$

2. Dynamical fidelity susceptibility χ at finite temperature

A possible generalization of the zero-temperature LE to finite temperatures is through the Uhlmann fidelity since the zero temperature $|\mathcal{A}(s)|$ is precisely the fidelity between the states $|\psi(0)\rangle$ and $\exp[-itH(s)]|\psi(0)\rangle$. Since the Uhlmann fidelity between two close mixed states is determined by the Bures metric, we begin by revisiting the derivation of the latter for the case of interest, i.e., two-level systems.

3. Bures metric for a two-level system

Take a curve of full-rank density operators $t\mapsto\rho(t)$ and a horizontal lift $t\mapsto W(t)$, with $W(0)=\sqrt{\rho(0)}$. Then, the Bures metric is given by

$$g_{\rho(t)}\left(\frac{d\rho}{dt},\frac{d\rho}{dt}\right) = \mathrm{Tr}\left\{\frac{dW^\dagger}{dt}\frac{dW}{dt}\right\}.$$

The horizontality condition is given by

$$W^\dagger\frac{dW}{dt} = \frac{dW^\dagger}{dt}W$$

for each t. In the full-rank case, we can find a unique Hermitian matrix $G(t)$, such that

$$\frac{dW}{dt} = G(t)W$$

solves the horizontality condition

$$W^\dagger\frac{dW}{dt} = W^\dagger GW = \frac{dW^\dagger}{dt}W.$$

Also, G is such that

$$\frac{d\rho}{dt} = \frac{d}{dt}(WW^\dagger) = G\rho + \rho G.$$

If L_ρ (R_ρ) is left (right) multiplication by ρ, we have, formally,

$$G = (L_\rho + R_\rho)^{-1}\frac{d\rho}{dt}.$$

Therefore,

$$g_{\rho(t)}\left(\frac{d\rho}{dt},\frac{d\rho}{dt}\right) = \mathrm{Tr}\left\{\frac{dW^\dagger}{dt}\frac{dW}{dt}\right\} = \mathrm{Tr}\{G^2\rho\} = \frac{1}{2}\mathrm{Tr}\{G(\rho G + G\rho)\} = \frac{1}{2}\mathrm{Tr}\left\{G\frac{d\rho}{dt}\right\} = \frac{1}{2}\mathrm{Tr}\left\{(L_\rho+R_\rho)^{-1}\frac{d\rho}{dt}\frac{d\rho}{dt}\right\}.$$

If we write $\rho(t)$ in the diagonal basis,

$$\rho(t) = \sum_i p_i(t)|i(t)\rangle\langle i(t)|$$

we find

$$\begin{aligned} g_{\rho(t)}&\left(\frac{d\rho}{dt},\frac{d\rho}{dt}\right) \\ &= \frac{1}{2}\mathrm{Tr}\left\{(L_\rho + R_\rho)^{-1}\frac{d\rho}{dt}\frac{d\rho}{dt}\right\} \\ &= \frac{1}{2}\sum_{i,j}\frac{1}{p_i(t)+p_j(t)}\langle i(t)|\frac{d\rho}{dt}|j(t)\rangle\langle j(t)|\frac{d\rho}{dt}|i(t)\rangle. \end{aligned}$$

Hence, using the diagonal basis of ρ, we can read off the metric tensor at ρ as

$$g_\rho = \frac{1}{2}\sum_{i,j}\frac{1}{p_i+p_j}\langle i|d\rho|j\rangle\langle j|d\rho|i\rangle.$$

This is the result for general full-rank density operators. For two-level systems, writing

$$\rho = \tfrac{1}{2}(1 - X^\mu\sigma_\mu),$$

and defining variables $|X| = r$ and $n^\mu = X^\mu/|X|$, we can express g_ρ as

$$\begin{aligned} g_\rho &= \left[\frac{1}{1+r}+\frac{1}{1-r}\right]d\rho_{11}^2 + d\rho_{12}d\rho_{21} \\ &= \frac{1}{1-r^2}d\rho_{11}^2 + d\rho_{12}d\rho_{21}, \end{aligned}$$

where we used $d\rho_{11} = -d\rho_{22}$. Notice that

$$d\rho_{11} = \tfrac{1}{2}\,\mathrm{Tr}\{d\rho U\sigma_3 U^{-1}\} = \tfrac{1}{2}\,\mathrm{Tr}\{d\rho n^\mu\sigma_\mu\},$$

where U is a unitary matrix diagonalizing ρ, $U\sigma_3U^{-1} = n^\mu\sigma_\mu$. Now,

$$d\rho = -\tfrac{1}{2}dX^\mu\sigma_\mu,$$

and hence

$$d\rho_{11} = -\tfrac{1}{2}dX^\mu n_\mu = -\tfrac{1}{2}dr.$$

On the other hand,

$$\begin{aligned} d\rho_{12}d\rho_{21} &= \tfrac{1}{4}\,\mathrm{Tr}\{d\rho U(\sigma_1 - i\sigma_2)U^{-1}\}\mathrm{Tr}\{d\rho U(\sigma_1 - i\sigma_2)U^{-1}\} \\ &= \tfrac{1}{4}[\mathrm{Tr}\{d\rho U\sigma_1U^{-1}\}\mathrm{Tr}\{d\rho U\sigma_1U^{-1}\} \\ &\quad + \mathrm{Tr}\{d\rho U\sigma_2U^{-1}\}\mathrm{Tr}\{d\rho U\sigma_2U^{-1}\}] \\ &= \tfrac{1}{4}\delta_{\mu\nu}(dX^\mu - n^\mu n_\lambda dX^\lambda)(dX^\nu - n^\nu n_\sigma dX^\sigma) \\ &= \tfrac{1}{4}r^2 dn^\mu dn_\mu, \end{aligned}$$

where we used the fact that the vectors (u,v) defined by the equations $U\sigma_1U^{-1} = u^\mu\sigma_\mu$ and $U\sigma_2U^{-1} = v^\mu\sigma_\mu$ form an orthonormal basis for the orthogonal complement in $\mathbb{R}^3$ of the line generated by n^μ (which corresponds to the tangent space to the unit sphere S^2 at n^μ). Thus, we obtain the final expression for the squared line element

$$ds^2 = \frac{1}{4}\left(\frac{dr^2}{1-r^2} + r^2\delta_{\mu\nu}dn^\mu dn^\nu\right). \tag{A2}$$

The above expression is ill defined for the pure state case of $r = 1$. Nevertheless, the limiting case of $r \to 1$ as the metric is smooth as we will now show by introducing another coordinate patch. The set of pure states is defined by $r = 1$, i.e., they correspond to the boundary of the three-dimensional ball $B = \{X \,:\, |X| = r \leqslant 1\}$ which, topologically, is the set of all density matrices in dimension 2. Introducing the change of variable $r = \cos u$, with $u \in [0,\pi/2)$, the metric becomes

$$ds^2 = \tfrac{1}{4}[du^2 + (\cos u)^2\delta_{\mu\nu}dn^\mu dn^\nu],$$

which is well defined also for the pure-state case of $r = \cos(0) = 1$. Restricting it to the unit sphere, the metric coincides with the Fubini-Study metric, also known as the quantum metric, on the space of pure states $\mathbb{C}P^1 \cong S^2$, i.e., the Bloch sphere. Therefore, there is no problem on taking the pure-state limit of this metric on the space of states since it reproduces the correct result.

4. Pullback of the Bures metric

We have a map

$$\begin{aligned} M \ni \lambda \mapsto \rho(\lambda) &= U(\lambda)\rho_0 U(\lambda)^{-1} \\ &= \tfrac{1}{2}U(\lambda)(I - X^\mu\sigma_\mu)U(\lambda)^{-1}, \end{aligned}$$

with

$$U(\lambda) = \exp[-itH(\lambda)],$$

and we take

$$\rho_0 = \frac{\exp[-\beta H(\lambda_0)]}{\mathrm{Tr}\{\exp[-\beta H(\lambda_0)]\}}, \text{ for some } \lambda_0 \in M.$$

We use the curve $[0,1] \ni s \mapsto \lambda(s)$, with $\lambda(0) = \lambda_0$, to obtain a curve of density operators

$$s \mapsto \rho(s) := \rho(\lambda(s)).$$

Notice that $\rho(0) = \rho_0$. The fidelity we consider is then

$$F(\rho(0),\rho(s)) = \mathrm{Tr}(\sqrt{\sqrt{\rho(0)}\rho(s)\sqrt{\rho(0)}}).$$

Recall that for 2×2 density operators of full rank the Bures line element reads as

$$ds^2 = \frac{1}{4}\left(\frac{dr^2}{1-r^2} + r^2\delta_{\mu\nu}dn^\mu dn^\nu\right),$$

where

$$n^\mu = X^\mu/|X| \text{ and } r = |X|,$$

with

$$\rho = \tfrac{1}{2}(I - X^\mu\sigma_\mu).$$

Now,

$$\rho(\lambda) = \tfrac{1}{2}[I - R^\mu_{\ \nu}(\lambda)X^\nu\sigma_\mu],$$

with $R^{\mu}_{\ \nu}(\lambda)$ being the unique SO(3) element satisfying

$$U(\lambda)\sigma_\mu U(\lambda)^{-1} = R^{\nu}_{\ \mu}(\lambda)\sigma_\nu.$$

We then have, pulling back the coordinates,

$$r(\lambda) = |X| = \text{const} \ \text{ and } \ n^\mu(\lambda) = R^{\mu}_{\ \nu}(\lambda)n^\nu.$$

Therefore,

$$\begin{aligned} ds^2 &= \frac{1}{4}r^2\delta_{\mu\nu}\frac{\partial n^\mu}{\partial\lambda^a}\frac{\partial n^\nu}{\partial\lambda^b}d\lambda^a d\lambda^b \\ &= \frac{1}{4}r^2\delta_{\mu\nu}n^\sigma n^\tau\frac{\partial R^{\mu}_{\ \sigma}}{\partial\lambda^a}\frac{\partial R^{\nu}_{\ \tau}}{\partial\lambda^b}d\lambda^a d\lambda^b, \end{aligned}$$

which in terms of the Euclidean metric on the tangent bundle of $\mathbb{R}^3$, denoted by $\langle *,*\rangle$, takes the form

$$g_{ab}(\lambda) = \frac{1}{4}r^2\left\langle R^{-1}\frac{\partial R}{\partial\lambda^a}n, R^{-1}\frac{\partial R}{\partial\lambda^b}n\right\rangle,$$

written in terms of the pullback of the Maurer-Cartan form in SO(3), $R^{-1}dR$. We can further pull back by the curve $s \mapsto \lambda(s)$ and evaluate at $s = 0$

$$\begin{aligned} \chi &:= g_{ab}(\lambda_0)\frac{\partial\lambda^a}{\partial s}(0)\frac{\partial\lambda^b}{\partial s}(0) \\ &= \frac{1}{4}r^2\left\langle R^{-1}\frac{\partial R}{\partial\lambda^a}n, R^{-1}\frac{\partial R}{\partial\lambda^b}n\right\rangle\frac{\partial\lambda^a}{\partial s}(0)\frac{\partial\lambda^b}{\partial s}(0), \end{aligned}$$

which gives us the expansion of the fidelity

$$F(s) \equiv F(\rho(0),\rho(s)) = 1 - \tfrac{1}{2}\chi s^2 + \cdots.$$

We now evaluate χ. Note that

$$\begin{aligned} &dU\sigma_\mu U^{-1} + U\sigma_\mu dU^{-1} \\ &\quad = U[U^{-1}dU,\sigma_\mu]U^{-1} = dR^{\nu}_{\ \mu}\sigma_\nu. \end{aligned}$$

Now, we can parametrize

$$U = y^0 I + iy^\mu\sigma_\mu, \text{ with } |y|^2 = 1.$$

Therefore,

$$\begin{aligned} U^{-1}dU &= (y^0 - iy^\mu\sigma_\mu)(dy^0 + idy^\nu\sigma_\nu) = i(y^0dy^\mu - y^\mu dy^0)\sigma_\mu + \frac{i}{2}(y^\mu dy^\nu - y^\nu dy^\mu)\varepsilon^{\lambda}_{\mu\nu}\sigma_\lambda; \\ [U^{-1}dU,\sigma_\kappa] &= -2\big[(y^0dy^\mu - y^\mu dy^0)\varepsilon^{\tau}_{\mu\kappa} + \tfrac{1}{2}(y^\mu dy^\nu - y^\nu dy^\mu)\varepsilon^{\lambda}_{\mu\nu}\varepsilon^{\tau}_{\lambda\kappa}\big]\sigma_\tau \\ &= -2\big[(y^0dy^\mu - y^\mu dy^0)\varepsilon^{\tau}_{\mu\kappa} + \tfrac{1}{2}(y^\mu dy^\nu - y^\nu dy^\mu)\big(\delta_{\mu\kappa}\delta^\tau_\nu - \delta^\tau_\mu\delta^\nu_\kappa\big)\big]\sigma_\tau \\ &= -2\big[(y^0dy^\mu - y^\mu dy^0)\varepsilon^{\tau}_{\mu\kappa} + (y^\kappa dy^\tau - y^\tau dy^\kappa)\big]\sigma_\tau \\ &= 2\big[(y^0dy^\mu - y^\mu dy^0)\varepsilon_{\mu\kappa}^{\ \tau} + (y^\tau dy^\kappa - y^\kappa dy^\tau)\big]\sigma_\tau \equiv (R^{-1}dR)^{\tau}_{\ \kappa}\sigma_\tau. \end{aligned}$$

Observe that for $H(\lambda) = x^\mu(\lambda)\sigma_\mu$, we have

$$y^0(\lambda) = \cos(|x(\lambda)|t) \text{ and } y^\mu = -\sin(|x(\lambda)|t)\frac{x^\mu(\lambda)}{|x(\lambda)|}.$$

Therefore,

$$\begin{aligned} dy^0 &= -\sin(|x(\lambda)|t)d|x(\lambda)|, \\ dy^\mu &= -\cos(|x(\lambda)|t)\frac{x^\mu(\lambda)}{|x(\lambda)|}d|x(\lambda)| - \sin(|x(\lambda)|t)d\left(\frac{x^\mu(\lambda)}{|x(\lambda)|}\right). \end{aligned}$$

After a bit of algebra, we get

$$\begin{aligned} y^0dy^\mu - y^\mu dy^0 &= \frac{x^\mu(\lambda)}{|x(\lambda)|}d|x(\lambda)| - \sin(|x(\lambda)|)\cos(|x(\lambda)|)d\left(\frac{x^\mu(\lambda)}{|x(\lambda)|}\right), \\ y^\mu dy^\nu - y^\nu dy^\mu &= 2\sin^2(|x(\lambda)|t)\frac{x^{[\mu}(\lambda)}{|x(\lambda)|}d\left(\frac{x^{\nu]}(\lambda)}{|x(\lambda)|}\right). \end{aligned}$$

Thus,

$$(R^{-1}dR)^{\tau}_{\kappa} = 2\frac{x^\mu(\lambda)}{|x(\lambda)|}d|x(\lambda)|\varepsilon_{\mu\kappa}^{\ \tau} - \sin(2|x(\lambda)|t)d\left(\frac{x^\mu(\lambda)}{|x(\lambda)|}\right)\varepsilon_{\mu\kappa}^{\ \tau} + 4\sin^2(|x(\lambda)|t)\frac{x^{[\tau}(\lambda)}{|x(\lambda)|}d\left(\frac{x^{\kappa]}(\lambda)}{|x(\lambda)|}\right).$$

At $s = 0$, $\lambda(0) = \lambda_0$ and the coordinate $n^\mu(\lambda(0)) = x^\mu(\lambda_0)/|x(\lambda_0)|$, so the previous expression reduces to

$$\begin{aligned} (R^{-1}dR(\lambda_0))^{\tau}_{\ \kappa}n^\kappa &= (R^{-1}dR(\lambda_0))^{\tau}_{\ \kappa}\frac{x^\kappa(\lambda_0)}{|x(\lambda_0)|} \\ &= -\sin(2|x(\lambda_0)|t)\frac{1}{|x(\lambda_0)|^2}\varepsilon^{\tau}_{\ \mu\kappa}x^\mu(\lambda_0)dx^\kappa(\lambda_0) - [1 - \cos(2|x(\lambda_0)|t)]d\left(\frac{x^\mu}{|x|}\right)(\lambda_0). \end{aligned}$$

Notice that the first term is perpendicular to the second. Therefore, we find

$$\begin{aligned}\chi ds^2 &= \frac{1}{4}r^2|R^{-1}dRn|^2 \\ &= \frac{1}{4}r^2\left[\sin^2(2|x(\lambda_0)|t)\frac{1}{|x(\lambda_0)|^4}\left(\delta^\lambda_\mu\delta^\sigma_\kappa - \delta^\sigma_\mu\delta^\lambda_\kappa\right)x^\mu(\lambda)dx^k(\lambda)x_\lambda(\lambda_0)dx_\sigma(\lambda_0) + [1-\cos(2|x(\lambda_0)|t)^2\langle Pdx(\lambda_0), Pdx(\lambda_0)\rangle]\right] \\ &= r^2\frac{\sin^2(|x(\lambda_0)|t)}{|x(\lambda_0)|^2}\left\langle P\frac{\partial x}{\partial\lambda^a}(\lambda_0), P\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle\frac{\partial\lambda^a}{\partial s}(0)\frac{\partial\lambda^b}{\partial s}(0)ds^2,\end{aligned}$$

where we have introduced the projector $P : T_x\mathbb{R}^3 = T_xS^2_{|x|} \oplus N_xS^2_{|x|} \to T_xS^2_{|x|}$ onto the tangent space of the sphere of radius $|x|$ at x. In other words, the pullback metric by ρ of the Bures metric at λ_0 is given by

$$g_{ab}(\lambda_0) = r^2\frac{\sin^2(|x(\lambda_0)|t)}{|x(\lambda_0)|^2}\left\langle P\frac{\partial x}{\partial\lambda^a}(\lambda_0), P\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle = \tanh^2(\beta|x(\lambda_0)|)\frac{\sin^2(|x(\lambda_0)|t)}{|x(\lambda_0)|^2}\left\langle P\frac{\partial x}{\partial\lambda^a}(\lambda_0), P\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle.$$

5. Dynamical interferometric susceptibility $\tilde{\chi}$ at finite temperature

We can replace the average $\langle e^{-itH(s)}\rangle \equiv \langle\psi(0)|e^{-itH(s)}|\psi(0)\rangle$ by the corresponding average of $e^{itH(0)}e^{-itH(s)}$ on the mixed state $\rho(\lambda_0) = \rho(0) = \exp[-\beta H(0)]/\mathrm{Tr}\{\exp[-\beta H(0)]\}$ (note its implicit temperature dependence):

$$\mathcal{A}(s) = \mathrm{Tr}\left(\rho(0)T\exp\left\{-i\int_0^t d\tau\, V(s,\tau)\right\}\right).$$

It is easy to see that $|\mathcal{A}(s)|^2$ has the same expansion as before with the average on $|\psi(0)\rangle$ replaced by the average on $\rho(0)$.

We now proceed to compute $\tilde{\chi}$, or equivalently $\tilde{g}_{ab}(\lambda_0)$, in the case of a two-level system, where we can write

$$\rho(\lambda) = \frac{e^{-\beta H(\lambda)}}{\mathrm{Tr}\{e^{-\beta H(\lambda)}\}} = \frac{1}{2}[I - X^\mu(\lambda)\sigma_\mu],$$

and define variables $r(\lambda) = |X(\lambda)|$ and $n^\mu(\lambda) = X^\mu(\lambda)/|X(\lambda)|$. Writing $H(\lambda) = x^\mu(\lambda)\sigma_\mu$ [and $H(s) \equiv H(\lambda(s))$], we have

$$V_a(\tau) = \frac{\partial x^\mu}{\partial\lambda^a}(\lambda_0)e^{i\tau H(0)}\sigma_\mu e^{-i\tau H(0)}.$$

Hence, its expectation value is

$$\langle V_a(\tau)\rangle = \frac{1}{\mathrm{Tr}\{e^{-\beta H(\lambda)}\}}\mathrm{Tr}\{e^{-\beta H(0)}\sigma_\mu\}\frac{\partial x^\mu}{\partial\lambda^a}(\lambda_0) = r(\lambda_0)n_\mu(\lambda_0)\frac{\partial x^\mu}{\partial\lambda^a}(\lambda_0) = X_\mu(\lambda_0)\frac{\partial x^\mu}{\partial\lambda^a}(\lambda_0),$$

which is independent of τ. We then have

$$\langle V_a(\tau_2)\rangle\langle V_b(\tau_1)\rangle = (r(\lambda_0))^2 n_\mu(\lambda_0)\frac{\partial x^\mu}{\partial\lambda^a}(\lambda_0)n_\nu(\lambda_0)\frac{\partial x^\nu}{\partial\lambda^b}(\lambda_0) = \tanh^2(\beta|x(\lambda_0)|)\frac{x_\mu}{|x(\lambda_0)|}\frac{\partial x^\mu}{\partial\lambda^a}(\lambda_0)\frac{x_\nu}{|x(\lambda_0)|}\frac{\partial x^\nu}{\partial\lambda^b}(\lambda_0),$$

where we used $X^\mu(\lambda_0) = \tanh(\beta|x(\lambda_0)|)x^\mu(\lambda_0)/|x(\lambda_0)|$. Now, using the cyclic property of the trace, we get

$$\begin{aligned}\frac{1}{2\,\mathrm{Tr}\{e^{-\beta H(0)}\}}\mathrm{Tr}\{e^{-\beta H(\lambda)}\{V_a(\tau_2), V_b(\tau_1)\}\} &= \frac{1}{2\,\mathrm{Tr}\{e^{-\beta H(0)}\}}\mathrm{Tr}\{e^{-\beta H(0)}\{\sigma_\mu,\sigma_\nu\}\}R^\mu_{\ \lambda}(\tau_2)R^\nu_{\ \sigma}(\tau_1)\frac{\partial x^\lambda}{\partial\lambda^a}(\lambda_0)\frac{\partial x^\sigma}{\partial\lambda^b}(\lambda_0) \\ &= \delta_{\mu\nu}R^\mu_{\ \lambda}(\tau_2)R^\nu_{\ \sigma}(\tau_1)\frac{\partial x^\lambda}{\partial\lambda^a}(\lambda_0)\frac{\partial x^\sigma}{\partial\lambda^b}(\lambda_0),\end{aligned} \tag{A3}$$

where $R^\mu_{\ \nu}(\tau)$ is the rotation matrix defined by the equation

$$e^{i\tau H(0)}\sigma_\nu e^{-i\tau H(0)} = R^\mu_{\ \nu}(\tau)\sigma_\mu. \tag{A4}$$

We can explicitly write $R^\mu_{\ \nu}(\tau)$ as

$$R^\mu_{\ \nu}(\tau) = \cos(2\tau|x(\lambda_0)|)\delta^\mu_\nu + [1-\cos(2\tau|x(\lambda_0)|)]n^\mu(\lambda_0)n_\nu(\lambda_0) + \sin(2\tau|x(\lambda_0)|)n^\lambda(\lambda_0)\varepsilon^\mu_{\lambda\nu}.$$

Using the previous equation, and because $\{R(\tau)\}$ forms a one-parameter group, we can write

$$\delta_{\mu\nu}R^\mu_{\ \lambda}(\tau_2)R^\nu_{\ \sigma}(\tau_1) = \delta_{\kappa\lambda}R^\kappa_{\ \sigma}(\tau_2-\tau_1).$$

Since $\tilde{\chi}$ [i.e., the metric $\tilde{g}_{ab}$; recall its zero-temperature expression (A1)] has to be symmetric under the label exchange $a \leftrightarrow b$, the relevant symmetric part of (A3) is

$$\cos[2(\tau_2-\tau_1)|x(\lambda_0)|]\delta_{\mu\nu}\frac{\partial x^\mu}{\partial\lambda^a}(\lambda_0)\frac{\partial x^\nu}{\partial\lambda^b}(\lambda_0) + (1-\cos[2(\tau_2-\tau_1)|x(\lambda_0)|])\frac{x_\mu}{|x(\lambda_0)|}\frac{\partial x^\mu}{\partial\lambda^a}(\lambda_0)\frac{x_\nu}{|x(\lambda_0)|}\frac{\partial x^\nu}{\partial\lambda^b}(\lambda_0).$$

Putting everything together gives

$$\left\langle \frac{1}{2}\{V_a(\tau_2),V_b(\tau_1)\}\right\rangle - \langle V_a(\tau_2)\rangle\langle V_b(\tau_1)\rangle = \cos\left[2(\tau_2-\tau_1)|x(\lambda_0)|\right]\left\langle P\frac{\partial x}{\partial\lambda^a}(\lambda_0), P\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle$$
$$+\left[1-\tanh^2(\beta|x(\lambda_0)|)\right]\left\langle \frac{x(\lambda_0)}{|x(\lambda_0)|},\frac{\partial x}{\partial\lambda^a}(\lambda_0)\right\rangle\left\langle \frac{x(\lambda_0)}{|x(\lambda_0)|},\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle.$$

The integral on τ_1 and τ_2 can now be performed, using

$$\int_0^t\int_0^t d\tau_2 d\tau_1 \cos[2(\tau_2-\tau_1)\varepsilon] = \int_0^t\int_0^t d\tau_2 d\tau_1[\cos(2\tau_2\epsilon)\cos(2\tau_1\epsilon)+\sin(2\tau_2\epsilon)\sin(2\tau_1\epsilon)]$$
$$= \frac{1}{4\epsilon^2}\left[\sin^2(2t\epsilon)+[\cos(2t\epsilon)-1][\cos(2t\epsilon)-1]\right]$$
$$= \frac{1}{4\epsilon^2}[2-2\cos(2t\epsilon)] = \frac{\sin^2(t\epsilon)}{\epsilon^2}.$$

So, the interferometric metric is

$$\tilde{g}_{ab}(\lambda_0) = \frac{\sin^2(|x(\lambda_0)|t)}{|x(\lambda_0)|^2}\left[\left\langle P\frac{\partial x}{\partial\lambda^a}(\lambda_0),P\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle\right] + t^2[1-\tanh^2(\beta|x(\lambda_0)|)]\left\langle \frac{x(\lambda_0)}{|x(\lambda_0)|},\frac{\partial x}{\partial\lambda^a}(\lambda_0)\right\rangle\left\langle \frac{x(\lambda_0)}{|x(\lambda_0)|},\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle.$$

The dynamical interferometric susceptibility is then given by

$$\tilde{\chi} = \tilde{g}_{ab}(\lambda_0)\frac{\partial\lambda^a}{\partial s}(0)\frac{\partial\lambda^b}{\partial s}(0) = \int_0^t\int_0^t d\tau_2 d\tau_1\left[\left\langle\frac{1}{2}\left\{\frac{\partial V}{\partial s}(0,\tau_2),\frac{\partial V}{\partial s}(0,\tau_1)\right\}\right\rangle - \left\langle\frac{\partial V}{\partial s}(0,\tau_2)\right\rangle\left\langle\frac{\partial V}{\partial s}(0,\tau_1)\right\rangle\right]$$
$$= \left\{\frac{\sin^2(|x(\lambda_0)|t)}{|x(\lambda_0)|^2}\left[\left\langle P\frac{\partial x}{\partial\lambda^a}(\lambda_0),P\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle\right]+t^2[1-\tanh^2(\beta|x(\lambda_0)|)]\left\langle\frac{x(\lambda_0)}{|x(\lambda_0)|},\frac{\partial x}{\partial\lambda^a}(\lambda_0)\right\rangle\left\langle\frac{x(\lambda_0)}{|x(\lambda_0)|},\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle\right\}\frac{\partial\lambda^a}{\partial s}(0)\frac{\partial\lambda^b}{\partial s}(0),$$

with the average taken with respect to the thermal state $\rho_0=\rho(0)\equiv\rho(\lambda_0)$. Also, as mentioned previously, we have the expansion

$$|\mathcal{A}(s)|^2 = |\mathrm{Tr}[\rho(0)\exp[itH(0)]\exp[-itH(s)]]|^2 = \left|\mathrm{Tr}\left[\rho(0)T\exp\left(-i\int_0^t d\tau V(s,\tau)\right)\right]\right|^2 = 1-\tilde{\chi}s^2+\cdots.$$

The difference between the two susceptibilities is given by

$$\tilde{\chi}-\chi = [1-\tanh^2(\beta|x(\lambda_0)|)]\left\{\frac{\sin^2(|x(\lambda_0)|t)}{|x(\lambda_0)|^2}\left[\left\langle P\frac{\partial x}{\partial\lambda^a}(\lambda_0),P\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle\right]\right.$$
$$\left.+t^2\left\langle\frac{x(\lambda_0)}{|x(\lambda_0)|},\frac{\partial x}{\partial\lambda^a}(\lambda_0)\right\rangle\left\langle\frac{x(\lambda_0)}{|x(\lambda_0)|},\frac{\partial x}{\partial\lambda^b}(\lambda_0)\right\rangle\right\}\frac{\partial\lambda^a}{\partial s}(0)\frac{\partial\lambda^b}{\partial s}(0).$$

As $\beta\to+\infty$, i.e., as the temperature goes to zero, the two susceptibilities are equal. Now, the function

$$f(t) = \frac{\sin^2(\epsilon t)}{\epsilon^2}$$

is well approximated by t^2 for small enough ϵ. In that case, the sum of the two terms appearing in the difference between susceptibilities is just proportional to the pullback Euclidean metric on $T\mathbb{R}^3$.

6. Pullback of the interferometric (Riemannian) metric on the space of unitaries

We first observe that each full-rank density operator ρ defines a Hermitian inner product in the vector space of linear maps of a Hilbert space $\mathscr{H}$, i.e., End($\mathscr{H}$), given by,

$$\langle A,B\rangle_\rho \equiv \mathrm{Tr}\{\rho A^\dagger B\}.$$

This inner product then defines a Riemannian metric on the trivial tangent bundle of the vector space End($\mathscr{H}$). Since the unitary group U($\mathscr{H}$) $\subset$ End($\mathscr{H}$), by restriction we get a Riemannian metric on U($\mathscr{H}$). If we choose ρ to be $e^{-\beta H(\lambda)}/\mathrm{Tr}\{e^{-\beta H(\lambda)}\}$, then take the pullback by the map $\Phi_t : M\ni\lambda_f\mapsto e^{-itH(\lambda_f)}\in\mathrm{U}(\mathscr{H})$ and evaluate at $\lambda_f=\lambda$, to obtain the desired metric.

Next, we show that this version of LE is closely related to the interferometric geometric phase introduced by Sjöqvist *et al.* [57,61]. To see this, consider the family of distances in U($\mathscr{H}$), d_ρ, parametrized by a full-rank density operator ρ, defined as

$$d_\rho^2(U_1,U_2) = \mathrm{Tr}\{\rho(U_1-U_2)^\dagger(U_1-U_2)\} = 2(1-\mathrm{Re}\langle U_1,U_2\rangle_\rho),$$

where $\langle *,* \rangle_\rho$ is the Hermitian inner product defined previously. In terms of the spectral representation of $\rho = \sum_j p_j |j\rangle\langle j|$, we have

$$\langle U_1, U_2 \rangle_\rho = \sum_j p_j \langle j | U_1^\dagger U_2 | j \rangle.$$

The Hermitian inner product is invariant under $U_i \mapsto U_i \cdot D$, $i = 1,2$, where D is a phase matrix

$$D = e^{i\alpha} \sum_j |j\rangle\langle j|.$$

For the interferometric geometric phase, one enlarges this gauge symmetry to the subgroup of unitaries preserving ρ, i.e., the gauge degree of freedom is $\mathrm{U}(1) \otimes \cdots \otimes \mathrm{U}(1)$. However, since we are interested in the interferometric LE previously defined, we choose not to do that, as we only need the diagonal subgroup, i.e., we only have a global phase. Next, promoting this global U(1)-gauge degree of freedom to a local one, i.e., demanding that we only care about unitaries modulo a phase, we see that, upon changing $U_i \mapsto U_i \cdot D_i$, $i = 1,2$, we have

$$\langle U_1, U_2 \rangle_\rho \mapsto \langle U_1 \cdot D_1, U_2 \cdot D_2 \rangle_\rho = \sum_j p_j \langle j | U_1^\dagger U_2 | j \rangle e^{i(\alpha_2 - \alpha_1)}.$$

We can choose gauges, i.e., D_1 and D_2, minimizing $d_\rho^2(U_1 \cdot D_1, U_2 \cdot D_2)$, obtaining

$$d_\rho^2(U_1 \cdot D_1, U_2 \cdot D_2) = 2(1 - |\langle U_1 \cdot D_1, U_2 \cdot D_2 \rangle_\rho|) = 2(1 - |\langle U_1, U_2 \rangle_\rho|).$$

Now, if $\{U_i = U(t_i)\}_{1 \leqslant i \leqslant N}$ were the discretization of a path of unitaries $t \mapsto U(t)$, $t \in [0,1]$, applying the minimization process locally, i.e., between adjacent unitaries U_{i+1} and U_i, in the limit $N \to \infty$ we get a notion of parallel transport on the principal bundle $\mathrm{U}(\mathscr{H}) \to \mathrm{U}(\mathscr{H})/\mathrm{U}(1)$. In particular, the parallel transport condition reads as

$$\mathrm{Tr}\left\{ \rho U^\dagger(t) \frac{dU}{dt}(t) \right\} = 0, \text{ for all } t \in [0,1].$$

If we take $\rho = \exp[-\beta H(\lambda_i)]/\mathrm{Tr}\{e^{-\beta H(\lambda_i)}\}$, $U_1 = \exp[-itH(\lambda_i)]$ and $U_2 = \exp[-itH(\lambda_f)]$, then the interferometric LE is

$$\mathcal{L}(t, \beta; \lambda_f, \lambda_i) = |\langle U_1, U_2 \rangle_\rho| = \langle \widetilde{U}_1, \widetilde{U}_2 \rangle_\rho,$$

where $\widetilde{U}_i = U_i \cdot D_i$ ($i = 1,2$) correspond to representatives satisfying the discrete version of the parallel transport condition.

[1] S. Sachdev, *Quantum Phase Transitions* (Wiley, Hoboken, NJ, 2007).

[2] L. D. Landau *et al.*, On the theory of phase transitions, Zh. Eksp. Teor. Fiz. **7**, 19 (1937).

[3] X.-G. Wen, Topological orders in rigid states, Int. J. Mod. Phys. B **4**, 239 (1990).

[4] J. M. Kosterlitz and D. J. Thouless, Long range order and metastability in two dimensional solids and superfluids. (Application of dislocation theory), J. Phys. C: Solid State Phys. **5**, L124 (1972).

[5] J. M. Kosterlitz and D. J. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, J. Phys. C: Solid State Phys. **6**, 1181 (1973).

[6] F. Duncan and M. Haldane, Continuum dynamics of the 1-D heisenberg antiferromagnet: Identification with the O (3) nonlinear sigma model, Phys. Lett. A **93**, 464 (1983).

[7] F. D. M. Haldane, Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly", Phys. Rev. Lett. **61**, 2015 (1988).

[8] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Quantized Hall Conductance in a Two-Dimensional Periodic Potential, Phys. Rev. Lett. **49**, 405 (1982).

[9] M. Z. Hasan and C. L. Kane, Colloquium: Topological insulators, Rev. Mod. Phys. **82**, 3045 (2010).

[10] X.-L. Qi and S.-C. Zhang, Topological insulators and superconductors, Rev. Mod. Phys. **83**, 1057 (2011).

[11] B. A. Bernevig and T. L. Hughes, *Topological Insulators and Topological Superconductors* (Princeton University Press, Princeton, NJ, 2013).

[12] C.-E. Bardyn, L. Wawer, A. Altland, M. Fleischhauer, and S. Diehl, Probing the Topology of Density Matrices, Phys. Rev. X **8**, 011035 (2018).

[13] F. Grusdt, Topological order of mixed states in correlated quantum many-body systems, Phys. Rev. B **95**, 075106 (2017).

[14] B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Uhlmann Connection in Fermionic Systems Undergoing Phase Transitions, Phys. Rev. Lett. **119**, 015702 (2017).

[15] B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, Boltzmann-gibbs states in topological quantum walks and associated many-body systems: Fidelity and uhlmann parallel transport analysis of phase transitions, J. Phys. A: Math. Theor. **50**, 365302 (2017).

[16] O. Andersson, I. Bengtsson, M. Ericsson, and E. Sjöqvist, Geometric phases for mixed states of the kitaev chain, Philosophical Trans. R. Soc. London A **374**, 20150231 (2016).

[17] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Symmetry-protected topological phases at finite temperature, 2D Mater. **2**, 034006 (2015).

[18] E. P. L. Van Nieuwenburg and S. D. Huber, Classification of mixed-state topology in one dimension, Phys. Rev. B **90**, 075141 (2014).

[19] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Uhlmann Phase as a Topological Measure for One-Dimensional Fermion Systems, Phys. Rev. Lett. **112**, 130401 (2014).

[20] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Two-Dimensional Density-Matrix Topological Fermionic Phases: Topological Uhlmann Numbers, Phys. Rev. Lett. **113**, 076408 (2014).

[21] C. E. Bardyn, M. A. Baranov, C. V. Kraus, E. Rico, A. İmamoğlu, P. Zoller, and S. Diehl, Topology by dissipation, New J. Phys. **15**, 085001 (2013).

[22] A. Rivas, O. Viyuela, and M. A. Martin-Delgado, Density-matrix chern insulators: Finite-temperature generalization of topological insulators, Phys. Rev. B **88**, 155141 (2013).

[23] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, Thermal instability of protected end states in a one-dimensional topological insulator, Phys. Rev. B **86**, 155140 (2012).

[24] O. Viyuela, A. Rivas, S. Gasparinetti, A. Wallraff, S. Filipp, and M. A. Martin-Delgado, Observation of topological Uhlmann phases with superconducting qubits, npj Quantum Inf. **4**, 10 (2018).

[25] M. Heyl, A. Polkovnikov, and S. Kehrein, Dynamical Quantum Phase Transitions in the Transverse-Field Ising Model, Phys. Rev. Lett. **110**, 135704 (2013).

[26] C. Karrasch and D. Schuricht, Dynamical phase transitions after quenches in nonintegrable models, Phys. Rev. B **87**, 195104 (2013).

[27] F. Andraschko and J. Sirker, Dynamical quantum phase transitions and the loschmidt echo: A transfer matrix approach, Phys. Rev. B **89**, 125120 (2014).

[28] S. Vajna and B. Dóra, Disentangling dynamical phase transitions from equilibrium phase transitions, Phys. Rev. B **89**, 161105 (2014).

[29] M. Heyl, Scaling and Universality at Dynamical Quantum Phase Transitions, Phys. Rev. Lett. **115**, 140602 (2015).

[30] C. Karrasch and D. Schuricht, Dynamical quantum phase transitions in the quantum potts chain, Phys. Rev. B **95**, 075143 (2017).

[31] J. C. Halimeh and V. Zauner-Stauber, Dynamical phase diagram of quantum spin chains with long-range interactions, Phys. Rev. B **96**, 134427 (2017).

[32] V. Zauner-Stauber and J. C. Halimeh, Probing the anomalous dynamical phase in long-range quantum spin chains through fisher-zero lines, Phys. Rev. E **96**, 062118 (2017).

[33] I. Homrighausen, N. O. Abeling, V. Zauner-Stauber, and J. C. Halimeh, Anomalous dynamical phase in quantum spin chains with long-range interactions, Phys. Rev. B **96**, 104436 (2017).

[34] S. Vajna and B. Dóra, Topological classification of dynamical phase transitions, Phys. Rev. B **91**, 155127 (2015).

[35] M. Schmitt and S. Kehrein, Dynamical quantum phase transitions in the kitaev honeycomb model, Phys. Rev. B **92**, 075114 (2015).

[36] J. C. Budich and M. Heyl, Dynamical topological order parameters far from equilibrium, Phys. Rev. B **93**, 085416 (2016).

[37] Z. Huang and A. V. Balatsky, Dynamical Quantum Phase Transitions: Role of Topological Nodes in Wave Function Overlaps, Phys. Rev. Lett. **117**, 086802 (2016).

[38] R. Jafari and H. Johannesson, Loschmidt Echo Revivals: Critical and Noncritical, Phys. Rev. Lett. **118**, 015701 (2017).

[39] N. Sedlmayr, M. Fleischhauer, and J. Sirker, Fate of dynamical phase transitions at finite temperatures and in open systems, Phys. Rev. B **97**, 045147 (2018).

[40] P. Jurcevic, H. Shen, P. Hauke, C. Maier, T. Brydges, C. Hempel, B. P. Lanyon, M. Heyl, R. Blatt, and C. F. Roos, Direct Observation of Dynamical Quantum Phase Transitions in an Interacting Many-Body System, Phys. Rev. Lett. **119**, 080501 (2017).

[41] N. Fläschner, D. Vogel, M. Tarnowski, B. S. Rem, D.-S. Lühmann, M. Heyl, J. C. Budich, L. Mathey, K. Sengstock, and C. Weitenberg, Observation of dynamical vortices after quenches in a system with topology, Nat. Phys. **14**, 265 (2018).

[42] H. T. Quan, Z. Song, X. F. Liu, P. Zanardi, and C. P. Sun, Decay of Loschmidt Echo Enhanced by Quantum Criticality, Phys. Rev. Lett. **96**, 140604 (2006).

[43] P. Zanardi and N. Paunković, Ground state overlap and quantum phase transitions, Phys. Rev. E **74**, 031123 (2006).

[44] P. Zanardi, H. T. Quan, X. Wang, and C. P. Sun, Mixed-state fidelity and quantum criticality at finite temperature, Phys. Rev. A **75**, 032109 (2007).

[45] F. Pollmann, S. Mukerjee, A. G. Green, and J. E. Moore, Dynamics after a sweep through a quantum critical point, Phys. Rev. E **81**, 020101 (2010).

[46] N. T. Jacobson, L. C. Venuti, and P. Zanardi, Unitary equilibration after a quantum quench of a thermal state, Phys. Rev. A **84**, 022115 (2011).

[47] L. Campos Venuti and P. Zanardi, Unitary equilibrations: Probability distribution of the Loschmidt echo, Phys. Rev. A **81**, 022113 (2010).

[48] L. C. Venuti, N. T. Jacobson, S. Santra, and P. Zanardi, Exact Infinite-Time Statistics of the Loschmidt Echo for a Quantum Quench, Phys. Rev. Lett. **107**, 010403 (2011).

[49] M. Heyl and J. C. Budich, Dynamical topological quantum phase transitions for mixed states, Phys. Rev. B **96**, 180304 (2017).

[50] U. Bhattacharya, S. Bandyopadhyay, and A. Dutta, Mixed state dynamical quantum phase transitions, Phys. Rev. B **96**, 180303(R) (2017).

[51] J. Lang, B. Frank, and J. C. Halimeh, Concurrence of dynamical phase transitions at finite temperature in the fully connected transverse-field Ising model, arXiv:1712.02175.

[52] J. C. Halimeh, V. Zauner-Stauber, I. P. McCulloch, I. de Vega, U. Schollwöck, and M. Kastner, Prethermalization and persistent order in the absence of a thermal phase transition, Phys. Rev. B **95**, 024302 (2017).

[53] N. Paunković, P. D. Sacramento, P. Nogueira, V. R. Vieira, and V. K. Dugaev, Fidelity between partial states as a signature of quantum phase transitions, Phys. Rev. A **77**, 052302 (2008).

[54] P. Zanardi, L. C. Venuti, and P. Giorda, Bures metric over thermal state manifolds and quantum criticality, Phys. Rev. A **76**, 062318 (2007).

[55] D. F. Abasto, A. Hamma, and P. Zanardi, Fidelity analysis of topological quantum phase transitions, Phys. Rev. A **78**, 010301 (2008).

[56] J.-H. Zhao and H.-Q. Zhou, Singularities in ground-state fidelity and quantum phase transitions for the Kitaev model, Phys. Rev. B **80**, 014403 (2009).

[57] E. Sjöqvist, A. K. Pati, A. Ekert, J. S. Anandan, M. Ericsson, D. K. L. Oi, and V. Vedral, Geometric Phases for Mixed States in Interferometry, Phys. Rev. Lett. **85**, 2845 (2000).

[58] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Topological field theory of time-reversal invariant insulators, Phys. Rev. B **78**, 195424 (2008).

[59] D. Chruscinski and A. Jamiolkowski, *Geometric Phases in Classical and Quantum Mechanics*, Vol. 36 (Springer, Berlin, 2012).

[60] A. Uhlmann, Transition probability (fidelity) and its relatives, Found. Phys. **41**, 288 (2011).

[61] D. M. Tong, E. Sjöqvist, L. C. Kwek, and C. H. Oh, Kinematic Approach to the Mixed State Geometric Phase in Nonunitary Evolution, Phys. Rev. Lett. **93**, 080405 (2004).

[62] C. N. Yang and T. D. Lee, Statistical theory of equations of state and phase transitions. I. Theory of condensation, Phys. Rev. **87**, 404 (1952).

[63] T. D. Lee and C. N. Yang, Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model, Phys. Rev. **87**, 410 (1952).

[64] M. E. Fisher, *Lectures in Theoretical Physics* (Gordon and Breach, New York, 1965).

[65] R. K. Pathria, *Statistical Mechanics* (Butterworth-Heinemann, Oxford, 1996).

[66] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Mathematics in Science and Engineering (Academic, New York, 1976).

[67] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht, Models of wave-function collapse, underlying theories, and experimental tests, Rev. Mod. Phys. **85**, 471 (2013).

[68] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Solitons in Polyacetylene, Phys. Rev. Lett. **42**, 1698 (1979).

[69] M. J. Rice and E. J. Mele, Elementary Excitations of a Linearly Conjugated Diatomic Polymer, Phys. Rev. Lett. **49**, 1455 (1982).

[70] S. N. Kempkes, A. Quelle, and C. Morais Smith, Universalities of thermodynamic signatures in topological phases, Sci. Rep. **6**, 38530 (2016).

Physics Letters A ••• (••••) •••–•••

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Entanglement swapping via three-step quantum walk-like protocol

Xiao-Man Li [a], Ming Yang [a,*], Nikola Paunković [b,c,*], Da-Chuang Li [d,*], Zhuo-Liang Cao [d,a]

[a] School of Physics and Material Science, Anhui University, Hefei, 230601, China
[b] Instituto de Telecomunicações, Av. Rovisco Pais, 1049-001, Lisboa, Portugal
[c] Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001, Lisboa, Portugal
[d] School of Electronic and Information Engineering, Hefei Normal University, Hefei, 230601, China

ARTICLE INFO

Article history:
Received 14 November 2016
Received in revised form 13 October 2017
Accepted 16 October 2017
Available online xxxx
Communicated by A. Eisfeld

Keywords:
Entanglement swapping
Quantum walk
Polarization entangled state

ABSTRACT

We present an entanglement swapping process for unknown nonmaximally entangled photonic states, where the standard Bell-state measurement is replaced by a three-step quantum walk-like state discrimination process, i.e., the practically nontrivial coupling element of two photons is replaced by manipulating their trajectories, which will greatly enrich the dynamics of the coupling between photons in realizing quantum computation, and reduce the integration complexity of optical quantum processing. In addition, the output state can be maximally entangled, which allows for entanglement concentration as well.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

In quantum communication and quantum computation, quantum entanglement finds many significant applications, including quantum teleportation [1], quantum superdense coding [2], quantum cryptography [3], etc. Typically, only maximally entangled states (MES) can lead to the perfect implementation of the above protocols. But in real experiments, unavoidable decoherence of quantum system is a serious hindrance to the realization of quantum information processing and quantum computation. In general, it is inevitable that the degree of entanglement decreases with the channel length, leading to an effective non-maximally entangled state. Undoubtedly, the use of non-MES could lead to severe decrease in the efficiency and fidelity of a quantum communication protocol. Therefore, creation of a MES from non-MESs attracts considerable attention in the community. To circumvent this problem, Schmidt projection scheme and Procrustean scheme have been proposed [4]. Although entanglement swapping scheme was proposed for entangling two remote qubits without direct interaction between them [5,6], it can be regarded as an entanglement concentration method too [7].

In the standard entanglement swapping process, a Bell state measurement constitutes the main swapping mechanism [5]. But, the realization of a Bell-state measurement (BSM) is not an easy task in experiment, so efforts have been made to design the entanglement swapping schemes without BSM. For instance, several implementation schemes of the entanglement swapping without BSM have been proposed both in cavity QED systems [8,9] and in quantum dot systems [10]. Essentially, the entanglement swapping schemes with or without BSM both require the coupling interactions between two qubits at the intermediate location. The coupling interactions in the swapping scheme with BSM can lead to a full discrimination of the four Bell states, meanwhile the coupling interactions in the swapping scheme without BSM can only lead to a partial discrimination of the four Bell states.

Recently, it was shown that quantum walk (QW) [11] can be used to implement a generalized measurement, i.e. a positive operator value measure (POVM) [12], and furthermore, a generalized measurement has been realized in discriminating non-orthogonal quantum states by executing a properly engineered QW [13]. But, in these advances, only the non-orthogonal quantum state discrimination of a single qubit has been studied and realized via QW. If this QW based state discrimination process can be generalized to the two-qubit case, it can be used to implement entanglement swapping too. In this paper, we present a three-step QW-like state discrimination scheme for four non-orthogonal two-qubit states, and thus the entanglement swapping scheme for two unknown non-maximally entangled states. The output state of the swapping process is maximally entangled, which allows for entanglement concentration as well.

In addition, the coupling between two qubits is the core part of the entanglement swapping schemes. But the current existing cou-

* Corresponding authors.
E-mail addresses: mingyang@ahu.edu.cn (M. Yang), npaunkovic@gmail.com (N. Paunković), dachuangli@ustc.edu.cn (D.-C. Li).

https://doi.org/10.1016/j.physleta.2017.10.022
0375-9601/© 2017 Elsevier B.V. All rights reserved.

pling mechanisms for photonic qubits and matter qubits are not suitable for integration. The coupling mechanism for two photons in our entanglement swapping process is realized by manipulating the trajectories of the photons. Thus, it is very easy to implement and integrate, and the versatile site-dependent operations and the intersite trajectory manipulations will greatly enrich the dynamics that this process can produce. Since our protocol can formally be described as a two-particle three-step QW, this opens new possibilities for quantum computation using the existing optical implementations of QWs.

This paper is organized as follows. In Sec. 2 we briefly introduce the concepts of entanglement swapping for non-maximally entangled states. In Sec. 3 we present our entanglement swapping scheme. Sec. 4 summarizes our results.

2. Entanglement swapping for unknown non-maximally entangled states

Suppose there are two pairs of polarization-entangled photons (1, 2) and (3, 4) shared by three remote users Alice, Bob and Clare:

$$|\psi\rangle_{12} = a\,|HH\rangle_{12} + b\,|VV\rangle_{12}\,, \tag{1}$$

$$|\psi\rangle_{34} = a\,|HH\rangle_{34} + b\,|VV\rangle_{34}\,, \tag{2}$$

where a, b satisfy the normalization condition $|a|^2 + |b|^2 = 1$. Photons (1, 2) belong to Alice and Clare, respectively, and photons (3, 4) belong to Clare and Bob. Here, $|H\rangle$ ($|V\rangle$) denotes the horizontal (vertical) polarization state of the photons. Without loss of generality, we can assume that the superposition coefficients a and b are all real numbers. Initially, the state of the two photon pairs is in a product form, which can be written as

$$\begin{aligned}|\psi\rangle_{1234} &= |\psi\rangle_{12}\otimes|\psi\rangle_{34}\\ &= \sqrt{\frac{a^4+b^4}{2}}(|\psi\rangle^{+}_{14}|\psi\rangle^{1}_{23} + |\psi\rangle^{-}_{14}|\psi\rangle^{2}_{23})\\ &\quad + ab(|\varphi\rangle^{+}_{14}|\psi\rangle^{3}_{23} + |\varphi\rangle^{-}_{14}|\psi\rangle^{4}_{23}),\end{aligned} \tag{3}$$

where

$$|\psi\rangle^{\pm}_{14} = \frac{1}{\sqrt{2}}(|HH\rangle_{14} \pm |VV\rangle)_{14}, \tag{4}$$

$$|\varphi\rangle^{\pm}_{14} = \frac{1}{\sqrt{2}}(|HV\rangle_{14} \pm |VH\rangle_{14}), \tag{5}$$

$$|\psi\rangle^{1}_{23} = \frac{1}{\sqrt{a^4+b^4}}(a^2\,|HH\rangle_{23} + b^2\,|VV\rangle_{23}), \tag{6}$$

$$|\psi\rangle^{2}_{23} = \frac{1}{\sqrt{a^4+b^4}}(a^2\,|HH\rangle_{23} - b^2\,|VV\rangle_{23}), \tag{7}$$

$$|\psi\rangle^{3}_{23} = \frac{1}{\sqrt{2}}(|HV\rangle_{23} + |VH\rangle_{23}), \tag{8}$$

$$|\psi\rangle^{4}_{23} = \frac{1}{\sqrt{2}}(|HV\rangle_{23} - |VH\rangle_{23}). \tag{9}$$

From Eq. (3), we can see that, as long as Clare, who has access to photons 2 and 3 (as depicted in Fig. 1), can discriminate the four states in Eqs. (6)–(9), four maximally entangled states in Eqs. (4), (5) can be generated among the two remote users Alice and Bob. But the four states in Eqs. (6)–(9) are not orthogonal to each other, and they cannot be distinguished with unit probability. So generalized measurements (POVMs) must be introduced to discriminate these non-orthogonal states [14,15]. Because the states to be swapped are unknown,[1] the states in Eqs. (6), (7) are unknown for us too, and thus these two states cannot be distinguished. Nevertheless, the states in Eqs. (8), (9) are totally known for us, so, in the following section, we will design a three-step QW-like process to discriminate these two states among the four non-orthogonal quantum states in Eqs. (6)–(9).

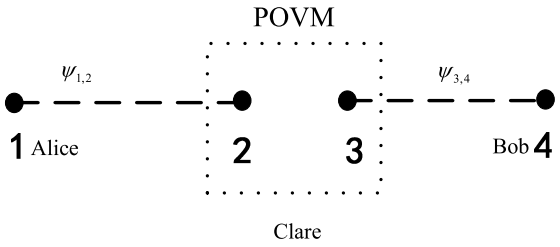


Fig. 1. The schematic diagram illustrating the procedure of entanglement swapping for unknown non-maximally entangled states.

3. Quantum walk-like swapping mechanism

In this section, we are going to design a three-step QW-like scheme to distinguish the two states in Eqs. (8), (9) from the four non-orthogonal quantum states in Eqs. (6)–(9), where the polarization degrees of photons 2, 3 are regarded as coin degrees of the three-step QW-like evolution, and the final position measurements on these two one-dimensional (1D) QW-like processes after three appropriately designed steps will tell us whether the discrimination succeeds or not.

Because our three-step scheme is a QW-like one, the state evolutions of the scheme are in similar forms as in QW systems, it is helpful for us to briefly review a standard 1D discrete-time QW [11]. The total Hilbert space of a walker consists of coin and position degrees of freedom, and is given by the tensor product $H \equiv H_{\mathcal{C}} \otimes H_{\mathcal{P}}$ of two subspaces spanned by $\{|H\rangle_{\mathcal{C}}, |V\rangle_{\mathcal{C}}\}$ and $\{|n\rangle_{\mathcal{P}}, n \in \mathbb{Z}\}$, respectively. Here, the subindex $\mathcal{C}$ denotes the coin degree, $\mathcal{P}$ denotes the position degree, and from now on they will be omitted for simplicity. One-step evolution of the system involves the coin flipping and conditional position shift based on the outcome of the coin flipping, and the corresponding unitary operation U is

$$U = S(C \otimes I), \tag{10}$$

where $C \in U(2)$ is the coin flipping operator, I is the identity operator in the position space, and the conditional position shift operator S takes the form $S = \sum_x (|x+1\rangle\langle x| \otimes |H\rangle\langle H| + |x-1\rangle\langle x| \otimes |V\rangle\langle V|)$. Without loss of generality, we assume the walker is at the position $x = 0$ initially, and the initial state of the coin is a superposition of $|H\rangle$ and $|V\rangle$ states. If the walk starts with the initial state $|\Psi(0)\rangle$, the final state of the system after t steps becomes

$$|\Psi(t)\rangle = U^t|\Psi(0)\rangle. \tag{11}$$

In our scheme, the states to be distinguished are two-photon (photons 2 and 3 at Clare's location) joint states in Eqs. (6)–(9) rather than the single-photon states, so the state evolutions for realizing this discrimination process are similar with the case of two walkers on two different lines. The joint Hilbert space of the two photons 2 and 3, on "line 2" and "line 3", respectively, is given by

$$H_{23} \equiv H_2 \otimes H_3 \equiv (H_{\mathcal{C}2} \otimes H_{\mathcal{P}2}) \otimes (H_{\mathcal{C}3} \otimes H_{\mathcal{P}3}). \tag{12}$$

[1] They are unknown in a sense that coefficients a and b are unknown, but the type of the states (superposition of both photon polarisations being either horizontal or vertical) is known.

Please cite this article in press as: X.-M. Li et al., Entanglement swapping via three-step quantum walk-like protocol, Phys. Lett. A (2017), https://doi.org/10.1016/j.physleta.2017.10.022

Here, H_2 and H_3 represent the Hilbert spaces of photons 2 and 3, respectively, each being isomorphic to the above Hilbert space H defined for the one-particle case. Note that, since the two lines are different, the photons are fully distinguishable by their spatial positions, i.e., the Hilbert space labels 2 and 3 are the physical labels associated to line 2 and 3, respectively, and thus the effects of particle statistics are not present. We will therefore for simplicity omit the symmetrization postulate and keep working in the full Hilbert space H, rather than in its symmetric subspace. The case when the two photons' spatial states overlap, making them indistinguishable, will be discussed later in the text. The one-particle coin degree of freedom is given by a photon's polarization (thus the notation $|H\rangle$ and $|V\rangle$, for horizontal and vertical polarization along given axes x and y, respectively). The corresponding unitary operation is given by

$$U_{23} = U_2 \otimes U_3, \tag{13}$$

where both U_2 and U_3 are isomorphic to the U defined above.

Assume both photons ("walkers") are at the position $x = 0$ initially, and the two polarization states ("coins") are initially ($t = 0$) prepared in a general (possibly entangled) state $|\Psi(0)\rangle_{\mathcal{C}_{23}} = \alpha|H\rangle_{\mathcal{C}2}|H\rangle_{\mathcal{C}3} + \beta|H\rangle_{\mathcal{C}2}|V\rangle_{\mathcal{C}3} + \gamma|V\rangle_{\mathcal{C}2}|H\rangle_{\mathcal{C}3} + \delta|V\rangle_{\mathcal{C}2}|V\rangle_{\mathcal{C}3}$ with the superposition coefficients satisfying the normalization condition. The overall composite system of the two photons is then in the following state (for simplicity, we write $|H,0\rangle_2$ to denote $|H\rangle_{\mathcal{C}2}|0\rangle_{\mathcal{P}2}$, etc.):

$$|\Psi(0)\rangle_{23} = \alpha|H,0\rangle_2|H,0\rangle_3 + \beta|H,0\rangle_2|V,0\rangle_3 + \gamma|V,0\rangle_2|H,0\rangle_3 + \delta|V,0\rangle_2|V,0\rangle_3. \tag{14}$$

After t-step evolution, the final state of the system becomes $|\Psi(t)\rangle_{23} = U_2^t \otimes U_3^t|\Psi(0)\rangle_{23}$. For properly engineered evolution, the photons with different initial polarization states will arrive at different positions. By projective measurements on the photons' positions, the initial polarization states can be discriminated.

During the swapping process described by Eq. (3), the initial polarization state of Clare's two photons is one of the four states in Eqs. (6)–(9), and the two photons both start from the origin position $x = 0$. Thus, the four possible initial position-polarization ("walker-coin") states are given by

$$|\psi(0)\rangle_{23}^1 = \frac{a^2|H,0\rangle_2|H,0\rangle_3 + b^2|V,0\rangle_2|V,0\rangle_3}{\sqrt{a^4+b^4}}, \tag{15}$$

$$|\psi(0)\rangle_{23}^2 = \frac{a^2|H,0\rangle_2|H,0\rangle_3 - b^2|V,0\rangle_2|V,0\rangle_3}{\sqrt{a^4+b^4}}, \tag{16}$$

$$|\psi(0)\rangle_{23}^3 = \frac{|H,0\rangle_2|V,0\rangle_3 + |V,0\rangle_2|H,0\rangle_3}{\sqrt{2}}, \tag{17}$$

$$|\psi(0)\rangle_{23}^4 = \frac{|H,0\rangle_2|V,0\rangle_3 - |V,0\rangle_2|H,0\rangle_3}{\sqrt{2}}. \tag{18}$$

The detailed three-step swapping mechanism is shown in Fig. 2. The polarization operators for each photon depend on the step and, for the first step, on the line as well (the polarization operator for line 2 is in the first step the identity I, while for line 3 the polarization operator is the NOT gate, with respect to the $\{|H\rangle, |V\rangle\}$ basis). The polarization operators are labelled by $C_{i,j}$, where the label $i = 1, 2, 3$ represents the step, while the label $j = 2, 3$ represents the line. Thus, in the first step the polarization operators are $C_{1,2} = I$ and

$$C_{1,3} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{19}$$

The polarization operators in the second step are also simple NOT gates,

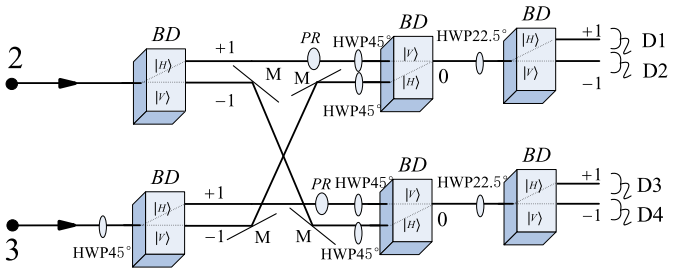


Fig. 2. Schematic drawing of the optical circuit for the three-step QW-like scheme that partially distinguishes the four possible two-photon polarization states. BD indicates beam displacers, by which vertically polarized photons are directly transmitted and horizontally polarized photons are moved up into a neighbouring mode with a lateral displacement. M denotes mirrors, and PR indicates phase retarder, which is used for phase compensation. HWP45° is a half-wave plate oriented at 45°, whose function can be expressed as $|H\rangle \to |V\rangle$ and $|V\rangle \to |H\rangle$. HWP22.5° is a half-wave plate oriented at 22.5°, which can induce the transformations $|H\rangle \to \frac{|H\rangle+|V\rangle}{\sqrt{2}}$ and $|V\rangle \to \frac{|H\rangle-|V\rangle}{\sqrt{2}}$. D_i $(i = 1, 2, 3, 4)$ are ideal photon detectors.

$$C_{2,2} = C_{2,3} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{20}$$

while in the third step polarization operators are Hadamard gates,

$$C_{3,2} = C_{3,3} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{21}$$

Each polarization rotation is followed by a conditional position shift operation S. The polarization operations can be implemented by half wave plates (HWPs) set at different orientations, and the conditional position shift operation S can be realized by birefringent calcite beam displacers (BDs).

Hence, at the end of the first step, the states are

$$|\psi(1)\rangle_{23}^1 = \frac{a^2|H,1\rangle_2|V,-1\rangle_3 + b^2|V,-1\rangle_2|H,1\rangle_3}{\sqrt{a^4+b^4}}, \tag{22}$$

$$|\psi(1)\rangle_{23}^2 = \frac{a^2|H,1\rangle_2|V,-1\rangle_3 - b^2|V,-1\rangle_2|H,1\rangle_3}{\sqrt{a^4+b^4}}, \tag{23}$$

$$|\psi(1)\rangle_{23}^3 = \frac{|H,1\rangle_2|H,1\rangle_3 + |V,-1\rangle_2|V,-1\rangle_3}{\sqrt{2}}, \tag{24}$$

$$|\psi(1)\rangle_{23}^4 = \frac{|H,1\rangle_2|H,1\rangle_3 - |V,-1\rangle_2|V,-1\rangle_3}{\sqrt{2}}. \tag{25}$$

After the first step, the two site-dependent ($x = -1$) paths of the two photons will be exchanged, which is the key part of this swapping mechanism and causes the coupling between the two photons. Note that this kind of coupling is simply done by redirecting and exchanging the paths of two photons, which is much simpler than the coupling induced by a beam splitter. As immediately after the exchange the two photons are still spatially distinguishable (in case both photons end up in line 2, the one that came from line 2 is now at position $+1$ while the one coming from line 3 is at position -1; in case both are at line 3, the roles are symmetric – the photon that stayed in line 3 is at position $+1$, while the other is at position -1), we can keep omitting symmetrization postulate. Nevertheless, as the position degrees of freedom of both photons are, after the exchange, enlarged, we will additionally label the position quantum numbers by n_2 and n_3 for the positions on line 2 and line 3, respectively, with $n_{2/3} \in \mathbb{Z}$. That is to say, the lines 2, 3 are exactly two different spatial locations, which enlarge the position space of the photon from the other line, for instance, line 2 is an enlargement of the position space of the photon in line 3 and vice versa. In other words, the relevant part of the "second photon" Hilbert space (the photon coming from the line 2) is now $H_2 = \text{span}\{|-1_2\rangle_2, |0_2\rangle_2, |+1_2\rangle_2\}, |-1_3\rangle_2, |0_3\rangle_2, |+1_3\rangle_2\}$,

and analogously for H_3. After the path exchange, the states evolve into

$$|\psi(1)\rangle_{23}^{1}=\frac{a^2|H,1_2\rangle_2|V,-1_2\rangle_3+b^2|V,-1_3\rangle_2|H,1_3\rangle_3}{\sqrt{a^4+b^4}}, \tag{26}$$

$$|\psi(1)\rangle_{23}^{2}=\frac{a^2|H,1_2\rangle_2|V,-1_2\rangle_3-b^2|V,-1_3\rangle_2|H,1_3\rangle_3}{\sqrt{a^4+b^4}}, \tag{27}$$

$$|\psi(1)\rangle_{23}^{3}=\frac{|H,1_2\rangle_2|H,1_3\rangle_3+|V,-1_3\rangle_2|V,-1_2\rangle_3}{\sqrt{2}}, \tag{28}$$

$$|\psi(1)\rangle_{23}^{4}=\frac{|H,1_2\rangle_2|H,1_3\rangle_3-|V,-1_3\rangle_2|V,-1_2\rangle_3}{\sqrt{2}}. \tag{29}$$

Note that for the first two states, $|\psi(1)\rangle_{23}^{1/2}$, the two photons end in the same line, while for the other two, $|\psi(1)\rangle_{23}^{3/4}$, they end up in different lines.

For the second step the polarization operators are expressed in Eq. (20). To make sure that when the two photons end up in the same line at the end of step 2 their spatial wave-functions fully overlap, we introduce two phase retarders (PRs), which are used for phase compensation. As their spatial wave-functions fully overlap, being two identical particles, the two photons would be indistinguishable after step 2, and the effects of quantum statistics would occur, i.e. the first two states $|\psi(2)\rangle_{23}^{1/2}$ would consist only of the "bunching" terms (a characteristic behaviour of indistinguishable bosons). Therefore, the four states after step 2 become (for simplicity, we now factor the overall polarization and position states; also, since the Hilbert space labels are now redundant, we omit them as well)

$$|\psi(2)\rangle^{1}=\frac{|H,V\rangle+|V,H\rangle}{\sqrt{2}}\otimes\frac{a^2|0_2,0_2\rangle+b^2|0_3,0_3\rangle}{\sqrt{a^4+b^4}}, \tag{30}$$

$$|\psi(2)\rangle^{2}=\frac{|H,V\rangle+|V,H\rangle}{\sqrt{2}}\otimes\frac{a^2|0_2,0_2\rangle-b^2|0_3,0_3\rangle}{\sqrt{a^4+b^4}}, \tag{31}$$

$$|\psi(2)\rangle^{3}=\frac{|V,V\rangle+|H,H\rangle}{\sqrt{2}}\otimes|0_2,0_3\rangle, \tag{32}$$

$$|\psi(2)\rangle^{4}=\frac{|V,V\rangle-|H,H\rangle}{\sqrt{2}}\otimes|0_2,0_3\rangle. \tag{33}$$

For the third step the polarization operators are expressed in Eq. (21). After the three-step evolution, the initial four possible position-polarization ("walker-coin") states become

$$|\psi(3)\rangle^{1}=\frac{1}{\sqrt{2}}\Big[|H,H\rangle\otimes\Big(\frac{a^2|1_2,1_2\rangle+b^2|1_3,1_3\rangle}{\sqrt{a^4+b^4}}\Big) -|V,V\rangle\otimes\Big(\frac{a^2|-1_2,-1_2\rangle+b^2|-1_3,-1_3\rangle}{\sqrt{a^4+b^4}}\Big)\Big], \tag{34}$$

$$|\psi(3)\rangle^{2}=\frac{1}{\sqrt{2}}\Big[|H,H\rangle\otimes\Big(\frac{a^2|1_2,1_2\rangle-b^2|1_3,1_3\rangle}{\sqrt{a^4+b^4}}\Big) -|V,V\rangle\otimes\Big(\frac{a^2|-1_2,-1_2\rangle-b^2|-1_3,-1_3\rangle}{\sqrt{a^4+b^4}}\Big)\Big], \tag{35}$$

$$|\psi(3)\rangle^{3}=\frac{1}{\sqrt{2}}\Big(|H,H\rangle\otimes|1_2,1_3\rangle+|V,V\rangle\otimes|-1_2,-1_3\rangle\Big), \tag{36}$$

$$|\psi(3)\rangle^{4}=\frac{-1}{\sqrt{2}}\Big(|H,V\rangle\otimes|1_2,-1_3\rangle+|V,H\rangle\otimes|-1_2,1_3\rangle\Big). \tag{37}$$

At the end of the third step, four photon detectors will be used to detect the position information of the two photons. From Eqs. (34)–(37), we can see that the position information of the two photons will help us to identify two polarization states in Eqs. (8), (9), i.e., the "successful" results are coming from the cases when each line ends up containing only one photon, while the "unsuccessful" ones are characterised by the fact that both photons end up in the same line. Specifically, if we observe a coincidence either between detectors $D1$ and $D3$ or $D2$ and $D4$, then the incident two-photon polarization state is $|\psi\rangle_{23}^{3}$. On the other hand, if we observe a coincidence between detectors $D1$ and $D4$ or $D2$ and $D3$, then the incident two-photon polarization state is $|\psi\rangle_{23}^{4}$. The other two incident polarization states will both lead only to single-detector clicks, such that either of the four detectors $D1$–$D4$ can click for both initial states $|\psi\rangle_{23}^{1}$ and $|\psi\rangle_{23}^{2}$, which correspond to an inconclusive result. The correspondence between the results of the coincidence measurements and the incident two-photon states is listed in Table 1.

Table 1
The results of the coincidence measurements versus the incident two-photon polarization states.

Initial polarization states	Clicks
$\|\psi\rangle_{23}^{1}$	D_1, or D_2, or D_3, or D_4
$\|\psi\rangle_{23}^{2}$	D_1, or D_2, or D_3, or D_4
$\|\psi\rangle_{23}^{3}$	D_1 and D_3, or D_2 and D_4
$\|\psi\rangle_{23}^{4}$	D_1 and D_4, or D_2 and D_3

After the identification of the state $|\psi\rangle_{23}^{3}$ or $|\psi\rangle_{23}^{4}$, the state of the two remote photons 1 and 4 is projected onto the maximally entangled state $|\varphi\rangle_{14}^{+}$ or $|\varphi\rangle_{14}^{-}$. That is to say, the entanglement swapping for unknown non-maximally entangled states is realized by two three-step QW-like operations, and thus the unknown non-maximally entangled states are concentrated into maximally entangled ones.

One may wonder if the swapping process might succeed when the spatial wave-functions of the two photons do not overlap at the end of step 2 (in general, this happens). Does the problem of temporal synchronicity of the two photons affect the final results? The following analysis shows that the problem of temporal synchronicity of the two photons is irrelevant to our swapping scheme. If the two photons from line 2 and line 3 who "meet" during the second shift operation have different times of arrival to the third BD, they would be distinguishable, and the results presented in (34) and (35) would be (note that now the Hilbert space labels 2 and 3 have physical meaning, denoting the different times of arrival of the two photons; also, for simplicity we omit the overall normalization $2\sqrt{a^4+b^4}$)

$$\begin{aligned}|\psi'(3)\rangle_{23}^{1}\propto a^2\Big(&|H,1_2\rangle_2|H,1_2\rangle_3+|H,1_2\rangle_2|V,-1_2\rangle_3\\&-|V,-1_2\rangle_2|H,1_2\rangle_3-|V,-1_2\rangle_2|V,-1_2\rangle_3\Big)\\&+b^2\Big(|H,1_3\rangle_2|H,1_3\rangle_3-|H,1_3\rangle_2|V,-1_3\rangle_3\\&+|V,-1_3\rangle_2|H,1_3\rangle_3-|V,-1_3\rangle_2|V,-1_3\rangle_3\Big),\end{aligned} \tag{38}$$

$$\begin{aligned}|\psi'(3)\rangle_{23}^{2}\propto a^2\Big(&|H,1_2\rangle_2|H,1_2\rangle_3+|H,1_2\rangle_2|V,-1_2\rangle_3\\&-|V,-1_2\rangle_2|H,1_2\rangle_3-|V,-1_2\rangle_2|V,-1_2\rangle_3\Big)\\&-b^2\Big(|H,1_3\rangle_2|H,1_3\rangle_3-|H,1_3\rangle_2|V,-1_3\rangle_3\\&+|V,-1_3\rangle_2|H,1_3\rangle_3-|V,-1_3\rangle_2|V,-1_3\rangle_3\Big),\end{aligned} \tag{39}$$

and the resulting states in Eqs. (36), (37) do not change. From the Eqs. (36)–(39), we can see that the state $|\psi\rangle_{23}^{3}$ or $|\psi\rangle_{23}^{4}$ still can be identified by the coincidence measurements listed in Table 1. That is to say, the fact that one detector from each line clicks, as opposed to the case when ones from only one line click, would

discriminate the "successful" swapping results from the "unsuccessful" ones. So the temporal asynchronization of the two photons does not alter the final result of the paper. Naturally, one may prefer the asynchronous case, because one can always delay one photon over the other, making the two photons fully distinguishable.

From the discussions in Refs. [13,16,17], we can see that our scheme is feasible and can be implemented experimentally. The two photons $2, 3$ can be injected into the two free-space modes 0 of two BDs. Subsequently they propagate through BDs, free-space modes ± 1 between BDs, and recombine at the second set of BDs, which forms an interferometric structure. After the second interferometric structure, the output photons will be coupled into single-mode fibers and subsequently detected by photon detectors (avalanche photodiodes). It was demonstrated that the interferometric structure used here is inherently stable, and no active phase locking is required [16], which demonstrates the feasibility of the current scheme too.

From the results of the scheme we can see that the 'click' or 'no-click' indicating the presence or the absence of photons in the corresponding mode is sufficient for us to confirm the success of the scheme, which means that only common single-photon detectors will be involved rather than the sophisticated photon-number-resolution detectors [18–20]. In addition, only two photons are involved in the swapping dynamics, so the two-photon coincidence measurements indicate and confirm the success of the swapping process. One may say that, in a real situation, a successful output may not induce a two-photon coincidence counting because of the non-perfect detection efficiency of photon detectors. Yes, it does happen sometimes, but this fact only decreases the success probability of the scheme without decreasing the fidelity of the output state. As long as a two-photon coincidence counting is registered, the swapping process succeeds. Currently, the detection efficiency of a photon detector can reach 93% [21], which only slightly decreases the success probability $2|ab|^2$ (see (3)) of our swapping scheme. On the other hand, the system error, caused by the imperfections of the optical components such as the dark counts of photon detectors, the nonplanar optical surfaces and the inaccurate angles of BDs and wave plates, will cause a imperfect fidelity 0.99 of the output state in our scheme [13,16,17].

4. Conclusion

In this paper, we demonstrated that much richer dynamics for two-photon states can be produced via the quantum walk-like evolutions, such as site-dependent operations and the intersite trajectory manipulations, which opens a new direction in realizing quantum computation. We proposed a three-step quantum walk-like entanglement swapping scheme for two unknown non-maximally entangled states. The swapping mechanism used here is realized by manipulating the trajectories of two photons, which greatly reduces the integration complexity. In addition, maximally entangled states are generated among two spatially separated particles from two unknown non-maximally entangled states, which means an entanglement concentration process can be realized via our scheme too.

The result presented in this paper is just a demonstration of a simple QW-like dynamics for two-photon states, and more new dynamics for quantum communication and quantum computation, such as the basic two-photon logic gates, will be our next study direction.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (NSFC) under Grants No. 11274010, No. 11204061, No. 11374085; Anhui Provincial Natural Science Foundation under Grant No. 1408085MA16; The Anhui Provincial Candidates for academic and technical leaders Foundation under Grant No. 2015H052. NP acknowledges the support of SQIG Security and Quantum Information Group and IT project QbigD funded by FCT PEst-OE/EEI/LA0008/2013 and UID/EEA/50008/2013.

References

[1] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, W.K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels, Phys. Rev. Lett. 70 (13) (1993) 1895–1899.
[2] C.H. Bennett, S.J. Wiesner, Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states, Phys. Rev. Lett. 69 (20) (1992) 2881–2884.
[3] A.K. Ekert, Quantum cryptography based on Bell's theorem, Phys. Rev. Lett. 67 (6) (1991) 661–663.
[4] C.H. Bennett, H.J. Bernstein, S. Popescu, B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A 53 (4) (1996) 2046–2052.
[5] M. Zukowski, A. Zeilinger, M.A. Horne, A.K. Ekert, 'Event-ready-detectors' Bell experiment via entanglement swapping, Phys. Rev. Lett. 71 (26) (1993) 4287–4290.
[6] J.W. Pan, D. Bouwmeester, H. Weinfurter, A. Zeilinger, Experimental entanglement swapping: entangling photons that never interacted, Phys. Rev. Lett. 80 (28) (1998) 3891–3894.
[7] S. Bose, V. Vedral, P.L. Knight, Purification via entanglement swapping and conserved entanglement, Phys. Rev. A 60 (1) (1999) 194–197.
[8] M. Yang, W. Song, Z.L. Cao, Entanglement swapping without joint measurement, Phys. Rev. A 71 (3) (2005) 034312.
[9] M. Yang, Y. Zhao, W. Song, Z.L. Cao, Entanglement concentration for unknown atomic entangled states via entanglement swapping, Phys. Rev. A 71 (4) (2005) 044302.
[10] X.J. Yuan, M. Yang, P. Dong, Z.L. Cao, Entanglement concentration of unknown states via the quantum dot plus double-sided cavity system, Laser Phys. Lett. 12 (11) (2015) 115204.
[11] Y. Aharonov, L. Davidovich, N. Zagury, Quantum random walks, Phys. Rev. A 48 (2) (1993) 1687–1690.
[12] P. Kurzyński, A. Wójcik, Quantum walk as a generalized measuring device, Phys. Rev. Lett. 110 (20) (2013) 200404.
[13] Z.H. Bian, J. Li, H. Qin, X. Zhan, R. Zhang, B.C. Sanders, P. Xue, Realization of single-qubit positive-operator-valued measurement via a one-dimensional photonic quantum walk, Phys. Rev. Lett. 114 (20) (2015) 203602.
[14] A. Chefles, Unambiguous discrimination between linearly-independent quantum states, Phys. Lett. A 239 (6) (1998) 339–347.
[15] A. Delgado, L. Roa, J.C. Retamal, C. Saavedra, Entanglement swapping via quantum state discrimination, Phys. Rev. A 71 (1) (2005) 012303.
[16] M.A. Broome, A. Fedrizzi, B.P. Lanyon, I. Kassal, A. Aspuru-Guzik, A.G. White, Discrete single-photon quantum walks with tunable decoherence, Phys. Rev. Lett. 104 (15) (2010) 153602.
[17] P. Xue, R. Zhang, H. Qin, X. Zhan, Z.H. Bian, J. Li, Barry C. Sanders, Experimental quantum-walk revival with a time-dependent coin, Phys. Rev. Lett. 114 (14) (2015) 140502.
[18] C.F. Wildfeuer, A.J. Pearlman, J. Chen, J. Fan, A. Migdall, J.P. Dowling, Resolution and sensitivity of a Fabry–Perot interferometer with a photon-number-resolving detector, Phys. Rev. A 80 (4) (2009) 043822.
[19] A.J. Miller, S.W. Nam, J.M. Martinis, A.V. Sergienko, Demonstration of a low-noise near-infrared photon counter with multiphoton discrimination, Appl. Phys. Lett. 83 (4) (2003) 791–793.
[20] A.E. Lita, A.J. Miller, S.W. Nam, Counting near-infrared single-photons with 95% efficiency, Opt. Express 16 (5) (2008) 3032–3040.
[21] F. Marsili, V.B. Verma, J.A. Stern, S. Harrington, A.E. Lita, T. Gerrits, I. Vayshenker, B. Baek, M.D. Shaw, R.P. Mirin, S.W. Nam, Detecting single infrared photons with 93% system efficiency, Nat. Photonics 7 (2013) 210–214.

Please cite this article in press as: X.-M. Li et al., Entanglement swapping via three-step quantum walk-like protocol, Phys. Lett. A (2017), https://doi.org/10.1016/j.physleta.2017.10.022

Journal of Physics A:
Mathematical and
Theoretical

PURPOSE-LED PUBLISHING™

PAPER

Boltzmann–Gibbs states in topological quantum walks and associated many-body systems: fidelity and Uhlmann parallel transport analysis of phase transitions

To cite this article: Bruno Mera *et al* 2017 *J. Phys. A: Math. Theor.* **50** 365302

View the article online for updates and enhancements.

You may also like

- Boltzmann-Gibbs entropy is sufficient but not necessary for the likelihood factorization required by Einstein
Constantino Tsallis and Hans J. Haubold
- Tsallis-thermometer: a QGP indicator for large and small collisional systems
Gábor Bíró, Gergely Gábor Barnaföldi and Tamás Sándor Biró
- The statistics of mesoscopic systems and the physical interpretation of extensive and non-extensive entropies
D V Anghel and A S Parvan

This content was downloaded from IP address 147.91.81.84 on 19/09/2025 at 14:40

IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **50** (2017) 365302 (17pp) https://doi.org/10.1088/1751-8121/aa820e

Boltzmann–Gibbs states in topological quantum walks and associated many-body systems: fidelity and Uhlmann parallel transport analysis of phase transitions

Bruno Mera[1,2,4], Chrysoula Vlachou[3,4], Nikola Paunković[3,4] and Vítor R Vieira[1,2]

[1] CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
[2] Departamento de Física, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
[3] Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
[4] Instituto de Telecomunicações, 1049-001 Lisbon, Portugal

E-mail: chrysoula.vlachou@tecnico.ulisboa.pt

Received 16 March 2017, revised 13 June 2017
Accepted for publication 25 July 2017
Published 11 August 2017



Abstract

We perform the fidelity analysis for Boltzmann–Gibbs-like states in order to investigate whether the topological order of 1D fermionic systems at zero temperature is maintained at finite temperatures. We use quantum walk protocols that are known to simulate topological phases and the respective quantum phase transitions for chiral symmetric Hamiltonians. Using the standard approaches of the fidelity analysis and the study of edge states, we conclude that no thermal-like phase transitions occur as temperature increases, i.e. the topological behaviour is washed out gradually. We also show that the behaviour of the Uhlmann geometric factor associated to the considered fidelity exhibits the same behaviour as the latter, thus confirming the results obtained using the previously established approaches.

Keywords: quantum walks, topological insulators, thermal states, fidelity, Uhlmann parallel transport, phase transitions

(Some figures may appear in colour only in the online journal)

 © 2017 IOP Publishing Ltd Printed in the UK

1. Introduction

Since the seminal paper of Haldane [1], where the anomalous Hall insulator was discovered, there has been an intense investigation on topological phases of matter. To characterise topological phases, invariants such as Chern numbers [2] and the Berry geometric phase [3], as well as non-local string parameters [4], were used as signatures of topological orders (see [5]). Regarding zero-temperature quantum phase transitions, the critical behaviour of systems featuring topological order is accompanied by the existence of edge states at the boundary between two distinct topological phases [6, 7]. Moving to open quantum systems, that are described by mixed, rather than pure, states, it is natural to ask if the topological order is still present, and investigate which are the appropriate quantities ('topological order parameters') to describe possible phase transitions. Various types of the mixed-state generalisations of geometric phases [8, 9] and closely related quantities [10–13], as well as the Chern values [14], were used to infer topological phase transitions of open systems (thermal and/or non-equilibrium).

In addition to the standard local symmetry breaking and the aforementioned global topological order parameters, several information-theoretic quantities were used to study phase transitions, such as entanglement measures [15–17] and the fidelity [18–22]. Being the measure of the distinguishability between two quantum states, the fidelity was used in numerous studies of both zero- and finite-temperature symmetry-breaking phase transitions, as well as of Kosterlitz–Thouless transitions [23, 24]. In [25] the authors analysed the intimate connection between the pure-state fidelity and the Berry phase, showing that the fidelity-induced Riemannian metric and the Berry curvature are the real and imaginary part, respectively, of the so-called quantum geometric tensor. Regarding mixed thermal states, the fidelity and the Uhlmann factor were studied for the case of the BCS superconductivity [26] and subsequently, in addition to BCS, for topological insulators and superconductors [27].

Recently, Kitagawa *et al* [28], showed that discrete-time quantum walks (QW) can realise topological phases in 1D and 2D for all the symmetry classes [29, 30] of free-fermion systems. In particular, they provide the QW protocols that simulate representatives of all topological phases, featured by the presence of robust symmetry-protected edge states (see also [31]). In general, QW realisations are particularly useful, because, in addition to the simplicity of their mathematical description, the parameters that define them can be easily controlled in the lab. The aforementioned topological quantum walks have been experimentally realised as periodically driven systems [32] and there are several experimental proposals for measuring topological invariants employing this approach [33–35].

We study topological features at finite temperatures for Boltzmann–Gibbs (BG) thermal-like states of single-particle 1D discrete-time topological QW and the corresponding many-body systems. The topological QW protocols we use simulate representatives of the chiral symmetric classes BDI and AIII of 1D topological insulators. We analyse the behaviour of the fidelity, the associated Uhlmann parallel transport and the edge states, and conclude that the effective temperature only smears out the topological features exhibited at zero temperature, without causing temperature-driven phase transitions.

This paper is organised as follows: in the next section, we describe the main topological features of QWs and their origin, and present the respective protocols. For a detailed and complete analysis see [28, 31]. In section 3 we present the BG states used: the single-particle QW states and their many-body counterparts. Furthermore, we clarify the relationship between them and explain the motivation for their use in different physical scenarios. In section 4, we present our results on the fidelity, the Uhlmann parallel transport and the edge states at finite temperatures and discuss the possibility of temperature-driven phase transitions. Finally, we summarise and discuss the results and point out possible directions of future work.

2. Topological quantum walks

We consider the *split-step* 1D discrete-time QW introduced in [28]. A single step is given by the unitary

$$U(\theta_1, \theta_2) = T_1 R(\theta_2) T_0 R(\theta_1), \tag{1}$$

where the shift operators are

$$T_c = \sum_x |x + (-1)^c\rangle \langle x| \otimes |c\rangle \langle c| + |x\rangle \langle x| \otimes |1 \oplus c\rangle \langle 1 \oplus c|, \tag{2}$$

with $c \in \{0, 1\}$ and $\oplus$ being addition modulo 2, and the coin operators are $R(\theta) = \mathrm{e}^{\mathrm{i}\frac{\theta}{2}\vec{\alpha}\cdot\vec{\sigma}}$, where $\vec{\alpha}$ is a unit 3D vector and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the Pauli vector.

As mentioned in the Introduction, topological QWs can be realised by means of periodically driven systems given by periodic time-dependent Hamiltonians $H(t + \delta t) = H(t)$, where δt represents the time of a single step. The evolution operator for one period of the driving, $[0, \delta t]$, called the Floquet operator, is given by

$$U(\delta t) = \mathcal{T}\mathrm{e}^{-\mathrm{i}\int_0^{\delta t} H(t)\mathrm{d}t}, \tag{3}$$

where $\mathcal{T}$ is the time ordering operator. Using homotopy theory, in [36] the authors propose a classification of periodically driven systems, according to the topological properties of their Floquet operators. They consider the Floquet operator in terms of a local effective Hamiltonian, given by $U(\delta t) = \mathrm{e}^{-\mathrm{i}H_{\mathrm{eff}}(\delta t)}$. They show that if $U(\delta t)$ is trivial under all homotopy groups, then the associated H_{eff} can exhibit non-trivial topological behaviour. The triviality of $U(\delta t)$ under the homotopy groups implies the existence of a gap in the spectrum of H_{eff} and if moreover H_{eff} has certain symmetries, the system supports the topological phases present in static topological insulators and superconductors, classified according to its dimension and the presence of these symmetries [29, 30].

The unitary operator that describes one step of the evolution of the split-step QW, as given in (1), is trivial under all homotopy groups, therefore we can define a local effective Hamiltonian

$$H_{\mathrm{eff}}(\theta_1, \theta_2) \equiv -\mathrm{i}\delta t^{-1} \log U(\theta_1, \theta_2), \tag{4}$$

whose quasienergy spectrum has a gap [36]. To fix the branch of the logarithm, we choose for energy spectrum the first Brillouin zone, as in [37, 38], obtaining the single-particle H_{eff} consistent with realistic many-body counterparts discussed in the next section. This way, the QW 'provides a stroboscopic simulation of the evolution generated by H_{eff} at the discrete times $N\delta t$' [28]. In other words, the evolution of the QW is performed in discrete time steps which last δt units of time each.

Due to translational invariance, we can write:

$$H_{\mathrm{eff}}(\theta_1, \theta_2) = \sum_k [E_k(\theta_1, \theta_2)\vec{n}_k(\theta_1, \theta_2) \cdot \vec{\sigma}] \otimes |k\rangle \langle k|, \tag{5}$$

where $\mathcal{B}$ is the first Brillouin zone, $E_k \geqslant 0$ and $\vec{n}_k$ are 3D unit vectors (for simplicity, we omit the θ-dependences). The eigenstates are given by

$$H_k |\pm\vec{n}_k\rangle = \pm E_k |\pm\vec{n}_k\rangle, \text{ with } H_k = E_k\vec{n}_k \cdot \vec{\sigma}, \tag{6}$$

forming the two energy bands $\{\pm E_k,\ k \in \mathcal{B}\}$. This type of Hamiltonians feature topological phases in the presence of time-reversal, particle-hole and/or chiral/sublattice symmetry [28, 36]. The symmetry classes we study in this paper (BDI and AIII) are chiral. This implies that $\vec{n}_k$, which defines the quantisation axis for each quasimomentum k, is restricted to lie

on a great circle of the Bloch sphere, determined by $\vec{\alpha}$ and θ_1. The number of times that $\vec{n}_k$ winds around the origin, as k ranges within the first Brillouin zone $\mathcal{B}$, is the *winding number* of the map between the two circles. This is exactly what gives rise to the topological features of the QW. The other symmetries are then determined by the choice of $\vec{\alpha}$ and/or by doubling the coin degree of freedom [28, 31]. For each symmetry class, we fix the value of θ_1 ($\theta_1 = -\pi/2$ for BDI and $\theta_1 = \pi/2$ for AIII) and vary θ_2 [28], so from now on, we consider $H_{\text{eff},\theta_1}(\theta) \equiv H_{\text{eff}}(\theta_1, \theta)$. By additionally varying θ along the line of the walk, it is possible to create a domain wall that separates two different phases, where the aforementioned edge states lie (see appendix A, for the details).

3. Boltzmann–Gibbs density operators

As described in the previous section, QWs can simulate topological phases of free-fermion systems at zero temperature. We wish to investigate if they can also be used to infer topological behaviour of many-body free-fermion systems at finite temperature also. To do that, we study two types of BG-like states for many-body systems, as well as their single-particle counterparts, for which we consider the effective Hamiltonian of the QW. Since topological QWs can be realised as periodically driven systems, it is natural to ask if such states can be considered in this context, since the energy is not conserved and the quasienergies (defined modulo 2π) which are conserved, have no natural ordering. It has been shown that these states, called Floquet-Gibbs states, can emerge under certain conditions [39, 40] (see also the justification for the existence of a 'quasienergy Brillouin zone' in the previous section). Following the discussion from the previous section, regardless of the realisations using periodically driven systems, the QWs that we consider can also be achieved using time-independent effective Hamiltonians (which were subject of a number of theoretical studies [28, 37, 38], and also experimentally realised [32, 41]), for which the BG states are well defined.

Let us consider a collection of fermion creation and annihilation operators $\{\psi_{k\sigma}, \psi^\dagger_{k\sigma} : k \in \mathcal{B},\ \sigma \in \{\uparrow, \downarrow\}\}$ and form the spinors $\Psi_k = (\psi_{k\uparrow}, \psi_{k\downarrow})^T$. The first state that we consider is the canonical ensemble given by,

$$\varrho^{(0)} = \frac{\mathrm{e}^{-\beta\mathcal{H}}}{\mathcal{Z}^{(0)}} = \frac{1}{\mathcal{Z}^{(0)}} \prod_{k\in\mathcal{B}} \exp(-\beta \Psi_k^\dagger H_k \Psi_k), \tag{7}$$

where $\mathcal{H}$ is given by the sum over the momenta of quadratic Hamiltonians $\mathcal{H} = \sum_k \Psi_k^\dagger H_k \Psi_k$, H_k is given by equation (6), and $\mathcal{Z}^{(0)} = \mathrm{tr}(\mathrm{e}^{-\beta\mathcal{H}})$ is the corresponding partition function (note that $\mathcal{H}$ conserves the particle number and its action on the whole Hilbert space is determined by the action on the single-particle sector). This state maximises the von Neumann entropy, subject to the constraint $\langle\mathcal{H}\rangle = \text{const}$.

The second one appears when one maximises the von Neumann entropy, subject to two constraints: the above mentioned energy constraint, as well as the constraints on the average number of particles $\langle n_k\rangle = \langle\Psi_k^\dagger \Psi_k\rangle$ which are constant in time, but in general different for each k. The state is of the form:

$$\varrho^{(1)} = \frac{\mathrm{e}^{-\beta\mathcal{O}}}{\mathcal{Z}^{(1)}} = \frac{1}{\mathcal{Z}^{(1)}} \prod_{k\in\mathcal{B}} \exp[-\beta(\Psi_k^\dagger H_k \Psi_k - \mu_k \Psi_k^\dagger \Psi_k)], \tag{8}$$

where $\mathcal{O} = \sum_k \mathcal{O}_k = \sum_k \mathcal{H}_k - \mu_k n_k$, with $\mu_k = -(1/\beta)\log(Z_k)$ being a momentum dependent chemical potential (which, in this case, coincides with the Helmholtz free energy associated to momentum k), and $\mathcal{Z}^{(1)} = \mathrm{tr}(\mathrm{e}^{-\beta\mathcal{O}})$ is the corresponding partition function. For

details concerning the derivation of these states or, more generally, states maximising the von Neumann entropy subject to constraints, see the appendix of [42]. In the field of quantum integrable systems, the state $\varrho^{(1)}$ is known as the 'generalised Gibbs ensemble', which was introduced in [43, 44] (for a review, see [45]).

Since the previous zero-temperature studies (both theoretical and experimental) of symmetry-protected topological orders were done in terms of single-particle quantum walks, we present the corresponding single-particle projected states $\rho^{(0)}$ and $\rho^{(1)}$ of the many-body counterparts (7) and (8), respectively.

The first is the standard thermal state, resulting from the effective Hamiltonian of the QW:

$$\rho^{(0)} = \frac{\mathrm{e}^{-\beta H_{\mathrm{eff}}}}{Z} = \frac{1}{Z}\sum_{k\in\mathcal{B}} \mathrm{e}^{-\beta H_k} \otimes |k\rangle\langle k|\,, \tag{9}$$

where $Z = \mathrm{tr}\mathrm{e}^{-\beta H_{\mathrm{eff}}}$, while the second is:

$$\rho^{(1)} = \frac{1}{\Omega}\sum_{k\in\mathcal{B}} \frac{\mathrm{e}^{-\beta H_k}}{Z_k} \otimes |k\rangle\langle k| = \frac{1}{\Omega}\sum_{k\in\mathcal{B}} \rho_k \otimes |k\rangle\langle k|\,, \tag{10}$$

where $Z_k = \mathrm{tr}\mathrm{e}^{-\beta H_k}$ and $\Omega = \sum_k 1$ is the k-space volume. Note that by tracing out the momenta, the state $\rho^{(0)}$ remains to be of the BG form (it is 'globally', with respect to k, thermal-like), while $\rho^{(1)}$ is not—only by measuring the momenta, the state collapses to a 'local' BG form ρ_k. Notice also that $Z = \sum_k Z_k$. The difference between $\rho^{(0)}$ and $\rho^{(1)}$ becomes even more clear by looking at their asymptotic behaviours. Namely, when $\beta \to +\infty$

$$\rho^{(0)} \to \frac{1}{|M|}\sum_{k\in M} |-\vec{n}_k\rangle\langle -\vec{n}_k| \otimes |k\rangle\langle k|\,, \tag{11}$$

where $M = \{k_* \in \mathcal{B} : E(k_*) = \max_{k\in\mathcal{B}} E(k)\}$ is the set of momenta minimising the lower band dispersion and $|-\vec{n}_k\rangle$ is defined in equation (6), while

$$\rho^{(1)} \to \frac{1}{\Omega}\sum_{k\in\mathcal{B}} |-\vec{n}_k\rangle\langle -\vec{n}_k| \otimes |k\rangle\langle k|\,, \tag{12}$$

is a statistical mixture of the *entire* lower band of the Hamiltonian.

Note that there exists a bijection between single-particle and quadratic many-body Hamiltonians, given by $H_k \leftrightarrow \mathcal{H}_k = \Psi_k^\dagger H_k \Psi_k$, where the left arrow represents the aforementioned projection onto the single-particle sector, thus inducing the corresponding bijections between the BG states $\rho^{(0)} \leftrightarrow \varrho^{(0)}$ and $\rho^{(1)} \leftrightarrow \varrho^{(1)}$.

4. The results: fidelity, Uhlmann parallel transport and the edge states

As mentioned in the Introduction, the fidelity is a measure of the distinguishability between two quantum states ρ and ρ', and is given as $F(\rho,\rho') \equiv \mathrm{tr}\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}$. By definition $F(\rho,\rho') \geqslant 0$, with the equality holding in the case of two completely distinguishable states, and $F(\rho,\rho') = 1$ if and only if $\rho = \rho'$, which is the maximum value fidelity can take in the case of two completely indistinguishable states. When a system undergoes a phase transition, the density matrix describing it changes drastically, resulting in a sudden drop of the fidelity. This is the main motivation for using fidelity to detect phase transitions. Furthermore, the fidelity is closely related to the Uhlmann connection via the Bures metric over the parameter space (for details, see [27]). In particular, the 'Uhlmann factor' U, given by $\sqrt{\rho}\sqrt{\rho'} = |\sqrt{\rho}\sqrt{\rho'}|U$,

quantifies the difference between the eigenbases of the two states ρ and ρ', and thus the distinguishability between them [26]. To evaluate the departure of U from identity we consider the quantity

$$\Delta(\rho,\rho') = F(\rho,\rho') - \mathrm{tr}(\sqrt{\rho}\sqrt{\rho'}) = \mathrm{tr}(|\sqrt{\rho}\sqrt{\rho'}|(I-U)), \tag{13}$$

which we also use for the detection of phase transitions.

In our analysis, we study the overall quantum states over the parameter space denoted by $\vec{q} = (\beta^{-1}, \theta)$, including temperature and the parameter of the Hamiltonian that drives the topological phase transition. We consider the fidelity F and the quantity Δ between two states ρ and ρ' separated in the parameters' space by an 'infinitesimal' displacement: namely $F(\rho,\rho')$ and $\Delta(\rho,\rho')$, where prime denotes the 'infinitesimally' close parameters. We can consider three different cases. In the first case, which is the most general, we probe the system with respect to both the parameter θ and the temperature T on the same time, that is:

$$\rho' = \rho(\vec{q}\prime) = \rho(\vec{q} + \delta\vec{q}),$$

with $\delta\vec{q} = (\delta\theta, \delta T)$ and $\|\delta\vec{q}\| << \|\vec{q}\|$, since $|\delta\theta| << |\theta|$ and $|\delta T| << |T|$. The analytic derivation of the expressions for F and Δ of the BG states considered can be found in appendix B. The other two cases occur when we wish to probe the system with respect to the parameter of the Hamiltonian and the temperature separately, these are for $\delta\vec{q} = (\delta\theta, 0)$ and $\delta\vec{q} = (0, \delta T)$, respectively. The quantitative results for the representative of the class BDI are given in figure 1. The results for the representative of the class AIII are qualitatively similar and, thus, we present them in appendix C. For both classes, the quantitative results that we present in this work are obtained in the most general case, where we probe the system with respect to θ and T on the same time, i.e. $\delta\vec{q} = (\delta\theta, \delta T) = (0.01, 0.01)$. In what follows we also comment on the results obtained for the other two cases $\delta\vec{q} = (\delta\theta, 0) = (0.01, 0)$ and $\delta\vec{q} = (0, \delta T) = (0, 0.01)$, however we do not present them, as they are qualitatively similar.

A unique feature of periodically driven systems and their corresponding effective Hamiltonians is that both energies $E_k = 0$ and $E_k = \pi$ correspond to a closed gap (the difference between the two energy levels $\pm E_k$ becomes zero modulo 2π). This special feature of QWs yields a surprising result in our analysis. In our study, we observe a different behaviour of the gap closing points with temperature, depending on whether they correspond to $E_k = 0$ (the two points with $\theta > 0$) or $E_k = \pi$ energy (the two points with $\theta < 0$). Whenever the gap closes, the vector $\vec{n}$ is ill-defined (for its specific form, see [28]), but the behaviour of F and Δ will be different due to their dependence on the entire energy spectrum. For the case of $E_k = 0$, they signal two isolated zero-temperature points of quantum phase transitions, corresponding to $\theta = \pi/2,\ 3\pi/2$ (in the case of Δ we notice small, but still present, peaks at these points). As the temperature increases, they are no longer signalling a phase transition and this is due to the dependence on the hyperbolic sine of E_k, which vanishes for $E_k = 0$ (see the respective formulae in appendix B), thus eliminating the $\vec{n}\cdot\vec{n}'$ term carrying the relevant topological features. The significance of these points of the zero-temperature quantum phase transitions on the system in the low-temperature regime, and the existence of possible crossovers, remain as open questions and require further investigation. In contrast to that, for the $E_k = \pi$ gap closing points, the phase transition lines survive with temperature, hence revealing a 'finite-temperature quantum phase transition' (a phase transition occurring at finite temperature, driven solely by the Hamiltonian's parameter(s), and not the temperature). Again, this can be understood through the dependence on E_k via hyperbolic functions which take finite values for $E_k = \pi$, thus maintaining the dependence on $\vec{n}\cdot\vec{n}'$.

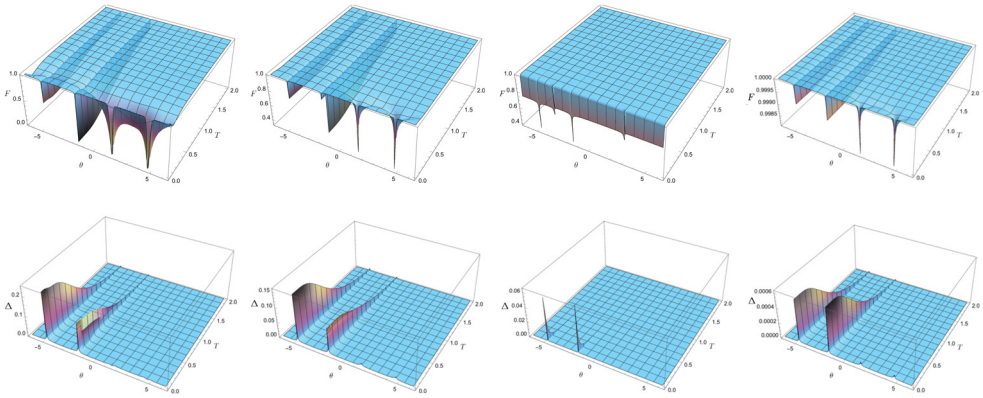


Figure 1. Fidelity (top) and $\Delta(\rho, \rho')$ (bottom) for the many-body states $\varrho^{(0)}$ and $\varrho^{(1)}$ and the single-particle states $\rho^{(0)}$ and $\rho^{(1)}$, for the BDI symmetry class. $\delta\theta = \theta' - \theta = 0.01$ and $\delta T = T' - T = 0.01$. The small step in the top middle-left plot is due to numerical instability (see appendix D).

Notice that for $\rho^{(0)}$ the qualitative behaviour is different from the other three types of states: the fidelity does not drop for $T = 0$ and $\theta = \pi/2, 3\pi/2$, while it does for two new $T = 0$ points at $\theta = \pm\pi$. The first difference is due to the fact that the zero temperature limit of $\rho^{(0)}$ projects only onto M given by the points of minimum energy $-E_k = -\pi$, see equation (11). Thus, both quantities do not see the critical momentum at which the gap closes at zero energy, which is above the lowest mode. The abrupt change of the fidelity at the points where $\theta = \pm\pi$ is the consequence of the enhanced zero-temperature ground state distinguishability due to the fact that E_k becomes constant and independent of k, i.e. the zero-temperature state projects onto the whole Brillouin zone ($M(\theta = \pm\pi) = \mathcal{B}$). Notice also that in the respective plot for Δ, the absence of peaks at $\theta = \pm\pi$ is consistent with the fact that the Uhlmann factor quantifies the change of eigenvectors, while the presence of two peaks at the corresponding fidelity plot is due to the flattening of the spectrum, i.e. solely to the change of eigenvalues.

Let us, now, comment on the results for the fidelity and the quantity Δ, in the case where we probe the system with respect to the parameter of the Hamiltonian and the temperature, separately. In the first case, where $\delta\theta = 0.01$ and $\delta T = 0$ ($\delta\vec{q} = (0.01, 0)$), the results are qualitatively the same as the ones presented in figure 1 for $\delta\vec{q} = (0.01, 0.01)$. In the second case, where we probe the system only with respect to temperature, that is $\delta\theta = 0, \delta T = 0.01$ or more concisely $\delta\vec{q} = (0, 0.01)$, the results for the fidelity are qualitatively the same as the ones for $\delta\vec{q} = (0.01, 0.01)$, while the results for the quantity Δ are different. In particular, we obtain that Δ is equal to zero everywhere. This triviality is due to the fact that Δ quantifies the difference between the eigenbases of ρ and ρ', as mentioned before and in this case that we only change the temperature and not the parameter of the Hamiltonian, the eigenbasis remains the same. On the other hand, the fidelity drops at the points of the phase transition, since it is also sensitive to changes in the spectrum.

We proceed to the results regarding edge states. The bulk-to-boundary principle predicts their existence on the boundary between distinct topological phases. These states are symmetry protected, i.e. they are robust against perturbations of the Hamiltonian which respect the symmetries of the system. In the case of pure states, QWs realise the aforementioned principle as shown in [28]. To this end, the authors introduce a spatial dependence on the parameter θ of the Hamiltonian in such a way that a phase boundary is created (see also appendix A for a

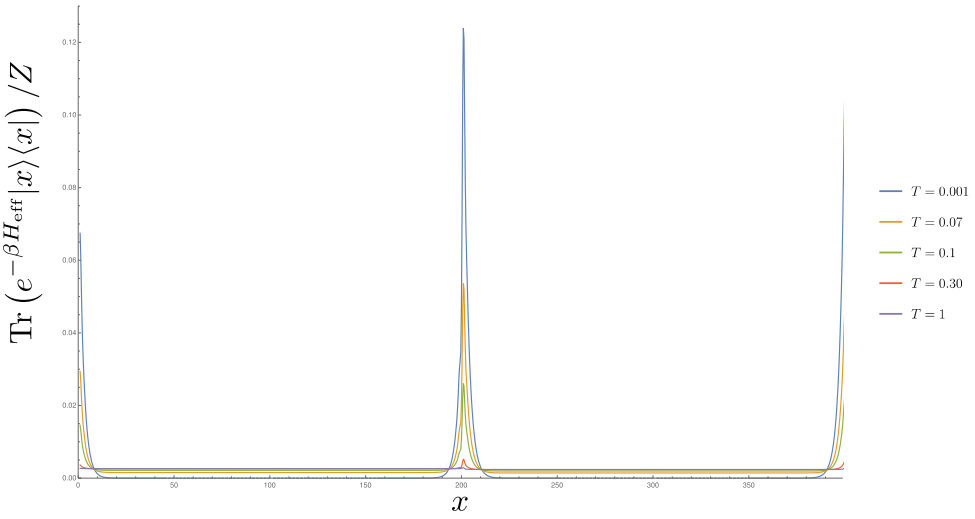


Figure 2. Position probability distribution of the QW as a function of the sites, $\mathrm{tr}(\mathrm{e}^{-H/T}\,|x\rangle\,\langle x|)/Z$. The Hamiltonian H is obtained by varying θ along x through a step-like function [28]. The domain wall is centred in the middle of the line. Periodic boundary conditions are taken, hence the edge state at the boundary.

brief description of the method). The edge states are then observed by evolving the walk *in time* from the initial state localised at the phase boundary: the probability to find the system in the initial position keeps being (considerably) higher than for the rest of the line. Instead, in our case, we are interested in probing the robustness of the edge states against temperature. Therefore, we study an ensemble of the BG type which corresponds to a stationary state of the split-step QW realising the BDI class representative, with the position-dependent coin operation, parameterised by its temperature β^{-1}. This stationary state appears naturally if one looks at the von Neumann evolution equation for the density matrix and imposes the stationarity condition. By diagonalising H_{eff} we obtain the localised states which can either have quasienergy 0 or π. Here, we choose a different order for the energy spectrum, by promoting the edge states with energies $E = \pi$ and $E = 0$ to be ground states, so that they survive at the zero-temperature limit.

At $T = 0$, the edge state has the major contribution to the probability distribution. As temperature increases, the edge states are smeared out, see figure 2. This was also observed in [14], by studying Chern values. Since the presence of an edge state is a clear manifestation of the topological order of a system at $T = 0$ then, the fact that it does not disappear for high temperatures, but it is rather smeared out, shows explicitly that the topological nature of the system is preserved in agreement with the fidelity analysis for the $E_k = \pi$ gap-closing points.

Our fidelity analysis reveals that for $T > 0$ there exist no thermal phase transitions (i.e. no temperature-driven phase transitions), as also shown in [27] for paradigmatic models of topological insulators and superconductors. However, here we observe finite temperature *parameter* θ-driven phase transitions, a property inherited from the temporal periodicity of the single-particle QW evolution. The non-existence of thermal phase transitions is further confirmed by the behaviour of the edge states, which are gradually smeared out as temperature increases, as also pointed out in [10]. Note that the edge states studied in the current paper are probed through a different method from the one used in [27]. There, a small chemical potential was introduced in such a way that the zero-temperature behaviour of the many-body density

operator is the projector onto the Fermi sea. Consequently, the existence of the edge states was signalled by the abrupt change in the number operator on the sites where they appeared.

5. Conclusions

We derived analytic expressions for the fidelity and the quantity Δ between two BG states for QW representatives of topological phases for the chiral symmetric classes BDI and AIII and their many-body counterparts. For the systems considered, the fidelity is detecting the points where the Bloch vector is ill-defined, which corresponds to the closure of the quasienergy gap. Since in our case the phase diagram is such that the closure of the gap always means the change from a trivial phase to a nontrivial one, and vice versa, the fidelity is capturing the topological phase transitions. Our results show the absence of temperature-driven phase transitions in two ways: through the fidelity analysis and through the behaviour of the edge states appearing on the phase boundary. In addition, the analysis of the fidelity and the Uhlmann connection, through the quantity Δ, shows the existence of finite-temperature transitions driven solely by the Hamiltonian's parameter θ, as a consequence of the temporal periodicity of the QW evolution. We would also like to point out that, providing that one of the goals of the paper is to study the finite-temperature behaviour of topological order in realistic many-body systems, the fact that the behaviour of the single-particle BG states is consistent with that of their many-body counterparts, shows that the analysis of the former can be a useful mathematical tool in the study of the latter.

Finally, we point out possible future lines of research. First, the same study could be applied to the rest of the symmetry classes in 1D and 2D using the protocols introduced in [28]. Further analysis of realistic noise effects that can give rise to BG states, in particular the single-particle ones, presents an interesting line of research (for a partial answer to this question, see also a recent study of the effects of thermal noise onto single-particle factor states of a topological insulator [14]). The above study could also be conducted for 2D QWs, as well as in the realm of multi-particle QWs.

The authors thank Sonja Barkhofen for providing clarifications regarding the experimental realisation presented in [41]. BM and CV acknowledge the support from DP-PMI and FCT (Portugal) through the grants SFRH/BD/52244/2013 and PD/BD/52652/2014, respectively. BM acknowledges the support of Physics of Information and Quantum Technologies Group. CV and NP acknowledge the support of SQIG–Security and Quantum Information Group and UID/EEA/50008/2013. NP acknowledges the IT project QbigD funded by FCT PEst-OE/EEI/LA0008/2013. VRV acknowledges support from FCT (Portugal) through Grant UID/CTM/04540/2013.

Appendix A. Quantum walk protocols

In this section, we briefly describe the quantum walk protocols we used for the simulation of the two chiral-symmetric classes BDI and AIII, as introduced in [28]. The Hilbert space of a one-dimensional discrete-time quantum walk (DTQW) is the tensor product of the position Hilbert space $H_p = \mathrm{span}\{|i\rangle_p \,|\, i \in \mathbb{Z}\}$, where i refer to the sites of an infinite line, with the two-dimensional coin Hilbert space $H_c = \mathrm{span}\{|0\rangle_c, |1\rangle_c\}$ (for simplicity, in the following we drop the subscripts p and c for the states). At each step, the walker is moving to the left or to the right on the line depending on its coin state. In a standard DTQW on a line, each step of the walk is given by the following unitary operator:

Table A1. Chiral Symmetric Classes in 1D and the respective quantum walk protocols. The values of the parameters θ_1 and θ_2 that correspond to distinct topological phases are shown, as well as the respective winding numbers ν of each phase.

Class	TRS	PHS	CS	Protocol	Parameters	ν
BDI	$\mathcal{T}^2 = 1$	$\mathcal{P}^2 = 1$	$(\Gamma^y_{\theta_1})^2 = 1$	$T_1 R_y(\theta_2) T_0 R_y(\theta_1)$	$\theta_1 = -\pi/2, \theta_2 = 3\pi/4$	$\nu = 0$
					$\theta_1 = -\pi/2, \theta_2 = \pi/4$	$\nu = 1$
AIII	Absent	Absent	$(\Gamma^\alpha_{\theta_1})^2 = 1$	$T_1 R_\alpha(\theta_2) T_0 R_\alpha(\theta_1)$	$\theta_1 = \pi/2, \theta_2 = 3\pi/4$	$\nu = 0$
					$\theta_1 = \pi/2, \theta_2 = \pi/4$	$\nu = 1$

$$U = T \cdot (I_p \otimes R), \tag{A.1}$$

where $R \in U(2)$ is a rotation in the 2-dimensional coin space, and T is the unitary operator describing the conditional shift of the walker (moving one position to the right when the coin state is $|0\rangle$, and to the left if it is $|1\rangle$):

$$T = \sum_x \left(|x+1\rangle\langle x| \otimes |0\rangle\langle 0| + |x-1\rangle\langle x| \otimes |1\rangle\langle 1| \right). \tag{A.2}$$

In [28] the authors show that this quantum walk simulates the Su–Schrieffer–Heeger (SSH) non-trivial topological phase with winding number 1. The SSH model is a representative of the BDI symmetry class, studied in the Letter. In order to study the behaviour of edge states, which appear on the boundary between two distinct topological phases, the authors in [28] introduce the so-called *split-step* quantum walk. As the name suggests, each step of the walk is split in two parts, each having the structure analogous to that of a standard DTQW,

$$U_{ss} = T_1 R(\theta_2) T_0 R(\theta_1), \tag{A.3}$$

where $R(\theta_i)$ represent rotation for the angle θ_i, along a fixed common axis $\vec{\alpha}$ (which, for simplicity, we omit to explicitly indicate), and

$$\begin{aligned} T_0 &= \sum_x \left(|x+1\rangle\langle x| \otimes |0\rangle\langle 0| + |x\rangle\langle x| \otimes |1\rangle\langle 1| \right), \\ T_1 &= \sum_x \left(|x-1\rangle\langle x| \otimes |1\rangle\langle 1| + |x\rangle\langle x| \otimes |0\rangle\langle 0| \right). \end{aligned} \tag{A.4}$$

For different values of the parameters θ_1 and θ_2 this protocol is shown to realise both trivial ($\nu = 0$) and non-trivial ($\nu = 1$) SSH topological phases. We choose θ_1 to be the same for both topological phases, while θ_2 is different. The closure of the gap implies a change of phase, in our case from a trivial to a non-trivial one and vice versa, so that by varying the value of θ_2, thus changing the energy spectrum, we are able to close the gap, and have a phase transition. That means that by making θ_2 change along the line, $\theta_2 = \theta_2(x)$, we can create a phase boundary on the site x_0 of the line, where θ_2 changes. Specifically, $\theta_2(x)$ is chosen to be $\theta_2(x) = \theta_2^{(t)} H(x_0 - x) + \theta_2^{(n)} H(x - x_0)$, where $H(x)$ is the Heaviside step function, and $\theta_2^{(t/n)}$ are the values of the angle for the trivial/non-trivial phase. Starting the quantum walk on this site (where the phase boundary is), permits us to detect the edge states, since even after many steps of the walk, there is still a large probability to find the walker at the origin, as a consequence of the overlap between the initial state with the edge state that is localised there [28, 31].

In order to simulate different symmetry classes (for the classification of topological insulators according to their symmetries, see [29, 30]), we further modify this basic split-step

protocol. By changing the spin rotation axis, we manage to break time-reversal (TRS), and particle-hole (PHS) symmetry, while maintaining chiral symmetry (CS), given by the operators $\mathcal{T}, \mathcal{P}$ and Γ, respectively. That leads us to a split-step protocol that simulates the symmetry class AIII.

The chiral symmetry operators are, for fixed θ_1, given as follows [28, 31]:

$$\begin{aligned}\Gamma^{y}_{\theta_1} &= \exp(-\mathrm{i}\pi\vec{A}^{y}_{\theta_1}\cdot\vec{\sigma}/2)\\ \Gamma^{\alpha}_{\theta_1} &= \exp(-\mathrm{i}\pi\vec{A}^{\alpha}_{\theta_1}\cdot\vec{\sigma}/2),\end{aligned} \tag{A.5}$$

for BDI and AIII, respectively. For each chiral class representative, this operator is determined by the vector $\vec{A}^{y/\alpha}_{\theta_1}$, perpendicular to unit vectors $y = [0, 1, 0]$ and $\alpha = 1/\sqrt{2}[0, 1, 1]$, and given by $\vec{A}^{y}_{\theta_1} = [\cos(\theta_1/2), 0, \sin(\theta_1/2)]$ (for BDI class) and $\vec{A}^{\alpha}_{\theta_1} = [\cos(\theta_1/2), -1/\sqrt{2}\sin(\theta_1/2), 1/\sqrt{2}\sin(\theta_1/2)]$ (for AIII class), respectively. Thus, operators $\Gamma^{y/\alpha}_{\theta_1}$ rotate $\vec{n}_{\theta_1}(k)$ by the angle π in a plane orthogonal to $\vec{A}^{y/\alpha}_{\theta_1}$. In class BDI, PHS is given by the operator

$$\mathcal{P} = \mathcal{K}, \tag{A.6}$$

where $\mathcal{K}$ denotes the complex conjugation operator, and TRS by the operator:

$$\mathcal{T} = \Gamma^{y}_{\theta_1}\mathcal{P}. \tag{A.7}$$

In table A1, we present the aforementioned 1D chiral classes, their symmetries, the quantum walk protocol that simulates them and the values of the parameters for each distinct topological phase, characterised by a winding number.

Appendix B. Analytic derivation of the closed expression for the fidelity

Given two thermal unnormalised states $\rho = \exp(-\beta H)$ and $\rho' = \exp(-\beta' H')$ the fidelity is given by

$$F(\rho, \rho') = \mathrm{Tr}\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}. \tag{B.1}$$

At the end of the calculation one must, of course, normalise the expression appropriately. In order to find closed expressions for the four possible density matrices we present in the main text, we will find e^C, such that

$$\mathrm{e}^A\mathrm{e}^B\mathrm{e}^A = \mathrm{e}^C, \tag{B.2}$$

for $A = -\beta H$, $B = -\beta' H'$ and, then we will take the square root of the result. Equivalently, the previous equation can be written as

$$\mathrm{e}^A\mathrm{e}^B = \mathrm{e}^C\mathrm{e}^{-A}. \tag{B.3}$$

By considering the Hamiltonians to be of the form $\vec{h}\cdot\vec{\sigma}$, we can write

$$\begin{aligned}\mathrm{e}^A &= a_0 + \vec{a}\cdot\vec{\sigma},\\ \mathrm{e}^B &= b_0 + \vec{b}\cdot\vec{\sigma},\\ \mathrm{e}^C &= c_0 + \vec{c}\cdot\vec{\sigma},\end{aligned} \tag{B.4}$$

where the coefficients are real and they satisfy the constraint:

$$1 = \det \mathrm{e}^A = a_0^2 - \vec{a}^2. \tag{B.5}$$

The same constraint holds for the determinants of B and C. We expand the LHS and the RHS of equation (B.3) and, by collecting the terms in 1, $\vec{\sigma}$ and $i\vec{\sigma}$, we get the following system of linear equations for c_0 and $\vec{c}$,

$$\begin{cases} a_0 b_0 + \vec{a}\cdot\vec{b} - a_0 c_0 + \vec{a}\cdot\vec{c} = 0 \\ a_0\vec{b} + b_0\vec{a} + \vec{a}c_0 - a_0\vec{c} = 0 \\ \vec{a}\times(\vec{b}-\vec{c}) = 0 \end{cases}. \tag{B.6}$$

The third equation has the solution $\vec{c} = \vec{b} + \lambda\vec{a}$, where λ is a real number. This means that the solution depends only on two real parameters: c_0 and λ. After some straightforward algebra, we obtain the solution:

$$\begin{bmatrix} c_0 \\ \lambda \end{bmatrix} = \begin{bmatrix} (2a_0^2 - 1)b_0 + 2a_0\vec{a}\cdot\vec{b} \\ 2(a_0 b_0 + \vec{a}\cdot\vec{b}) \end{bmatrix}. \tag{B.7}$$

Next, we set $A = -\beta H/2 \equiv -\xi\vec{x}\cdot\vec{\sigma}/2$ and $B = -\beta' H' \equiv -\zeta\vec{y}\cdot\vec{\sigma}$, with $\vec{x}^2 = \vec{y}^2 = 1$ and ξ, ζ are real parameters, meaning,

$$a_0 = \cosh(\xi/2) \text{ and } \vec{a} = -\sinh(\xi/2)\vec{x}, \tag{B.8}$$

$$b_0 = \cosh(\zeta) \text{ and } \vec{b} = -\sinh(\zeta)\vec{y}. \tag{B.9}$$

If we write $C = \rho\vec{z}\cdot\vec{\sigma}$ (which follows from the property of the product of special matrices) we get,

$$c_0 = \cosh(\rho) = \cosh(\xi)\cosh(\zeta) + \sinh(\xi)\sinh(\zeta)\vec{x}\cdot\vec{y}. \tag{B.10}$$

Using the formula $\cosh(\rho/2) = \sqrt{(1+\cosh(\rho))/2}$, we obtain,

$$\mathrm{tr}(\mathrm{e}^{C/2}) = 2\sqrt{\frac{(1+\cosh(\xi)\cosh(\zeta) + \sinh(\xi)\sinh(\zeta)\vec{x}\cdot\vec{y})}{2}}, \tag{B.11}$$

and by substituting, we end up with the following formula:

$$\mathrm{tr}(\sqrt{\mathrm{e}^{-\beta H/2}\mathrm{e}^{-\beta' H'}\mathrm{e}^{-\beta H/2}}) = 2\sqrt{\frac{(1+\cosh(\beta\Delta)\cosh(\beta'\Delta') + \sinh(\beta\Delta)\sinh(\beta'\Delta')\vec{n}\cdot\vec{n}')}{2}}. \tag{B.12}$$

To be able to compute all the fidelities, we will just need the following expression relating the traces of quadratic many-body fermion Hamiltonians (preserving the number operator) and the single-particle sector Hamiltonian obtained by projection:

$$\mathrm{tr}\mathrm{e}^{-\beta\mathcal{H}} = \mathrm{tr}\mathrm{e}^{-\beta\Psi^\dagger H\Psi} = \det(I + \mathrm{e}^{-\beta H}). \tag{B.13}$$

From the previous results, it is straightforward to derive the following formulae for the fidelities concerning all the *normalised* states discussed in the main text:

B.1. Single-body

$$\begin{aligned} F(\rho^{(0)}, \rho'^{(0)}) &= \frac{\sum_{k\in\mathcal{B}}\mathrm{tr}(\mathrm{e}^{-C_k/2})}{\sqrt{\sum_{k\in\mathcal{B}}\mathrm{tr}(\mathrm{e}^{-\beta H_k})\sum_{k\in\mathcal{B}}\mathrm{tr}(\mathrm{e}^{-\beta' H'_k})}} \\ &= \frac{\sum_{k\in\mathcal{B}}\sqrt{2\,(1+\cosh(E_k/2T)\cosh(E'_k/2T') + \sinh(E_k/2T)\sinh(E'_k/2T')\vec{n}_k\cdot\vec{n}'_k)}}{\sqrt{\sum_{k\in\mathcal{B}}2\cosh(E_k/2T)\sum_{k\in\mathcal{B}}2\cosh(E'_k/2T')}} \end{aligned} \tag{B.14}$$

$$F(\rho^{(1)},\rho'^{(1)}) = \sum_{k\in\mathcal{B}} \frac{\mathrm{tr}(\mathrm{e}^{-C_k/2})}{\sqrt{\mathrm{tr}(\mathrm{e}^{-\beta H_k})\mathrm{tr}(\mathrm{e}^{-\beta' H'_k})}}$$
$$= \sum_{k\in\mathcal{B}} \frac{\sqrt{2\,(1+\cosh(E_k/2T)\cosh(E'_k/2T') + \sinh(E_k/2T)\sinh(E'_k/2T')\vec{n}_k\cdot\vec{n}'_k)}}{\sqrt{2\cosh(E_k/2T)2\cosh(E'_k/2T')}} \tag{B.15}$$

B.2. Many-body

$$F(\varrho^{(0)},\varrho'^{(0)}) = \prod_{k\in\mathcal{B}} \frac{\mathrm{tr}(\mathrm{e}^{-\mathcal{C}_k/2})}{\mathrm{tr}(\mathrm{e}^{-(\beta\mathcal{H}_k)})\mathrm{tr}(\mathrm{e}^{-\beta'\mathcal{H}'_k})}$$
$$= \prod_{k\in\mathcal{B}} \frac{\det(I+\mathrm{e}^{-C_k/2})}{\det{}^{1/2}(I+\mathrm{e}^{-\beta H_k})\det{}^{1/2}(I+\mathrm{e}^{-\beta' H'_k})}$$
$$= \prod_{k\in\mathcal{B}} \frac{2+\sqrt{2\,(1+\cosh(E_k/2T)\cosh(E'_k/2T') + \sinh(E_k/2T)\sinh(E'_k/2T')\vec{n}_k\cdot\vec{n}'_k)}}{\sqrt{[2+2\cosh(E_k/2T)][2+2\cosh(E'_k/2T')]}} \tag{B.16}$$

$$F(\varrho^{(1)},\varrho'^{(1)}) = \prod_{k\in\mathcal{B}} \frac{\mathrm{tr}(\mathrm{e}^{-(\mathcal{C}_k-\log(Z_k)n_k-\log(Z'_k)n_k)/2})}{\mathrm{tr}(\mathrm{e}^{-(\beta\mathcal{H}_k-\log(Z_k)n_k)})\mathrm{tr}(\mathrm{e}^{-(\beta'\mathcal{H}'_k-\log(Z'_k)n_k)})}$$
$$= \prod_{k\in\mathcal{B}} \frac{\det(I+\mathrm{e}^{-C_k/2-\log(Z_k)/2-\log(Z'_k)/2})}{\det{}^{1/2}(I+\mathrm{e}^{-(\beta H_k-\log(Z_k))})\det{}^{1/2}(I+\mathrm{e}^{-(\beta' H'_k-\log(Z'_k))})}$$
$$= \prod_{k\in\mathcal{B}} \left(1+\frac{[2\,(1+\cosh(E_k/2T)\cosh(E'_k/2T') + \sinh(E_k/2T)\sinh(E'_k/2T')\vec{n}_k\cdot\vec{n}'_k)]}{\sqrt{(2\cosh(E_k/2T))(2\cosh(E'_k/2T'))}}\right.$$
$$\left. + \frac{1}{(2\cosh(E_k/2T))(2\cosh(E'_k/2T'))}\right) \times \left(\sqrt{[2+(2\cosh(E_k/2T))^{-2}][2+(2\cosh(E'_k/2T'))^{-2}]}\right)^{-1}, \tag{B.17}$$

where the single-particle partition function for momentum k, Z_k, is given by

$$Z_k = \mathrm{tr}\mathrm{e}^{-H_k/T} = 2\cosh(E_k/2T), \tag{B.18}$$

the matrix C_k is such that $\mathrm{e}^{-C_k} = \mathrm{e}^{-\beta H_k/2}\mathrm{e}^{-\beta' H'_k}\mathrm{e}^{-\beta H_k/2}$ and $\mathcal{C}_k = \Psi_k^\dagger C_k \Psi_k$ is the corresponding many-body quadratic operator.

To compute the quantity $\Delta(\rho,\rho')$ we need, in addition, $\mathrm{tr}\sqrt{\rho}\sqrt{\rho'}$. The derivation is similar but simpler, hence we shall omit the proof.

Appendix C. Results for the representative of AIII symmetry class

In figure C1 we present the quantitative results for the fidelity and the quantity Δ, in the case of the AIII symmetry class representative. Here, we also observe the different behaviour of the gap closing points with temperature, depending on whether they correspond to $E_k = 0$ (in this case, these are the two points with $\theta < 0$) or $E_k = \pi$ energy (the two points with $\theta > 0$). While the results for the representatives of BDI and AIII classes are qualitatively similar, the zero-temperature fidelity corresponding to the $\rho^{(0)}$ states for the AIII representative only exhibits drops in the gap closing points ($-E_k = -\pi$), as the spectrum is always non-trivially k-dependent.

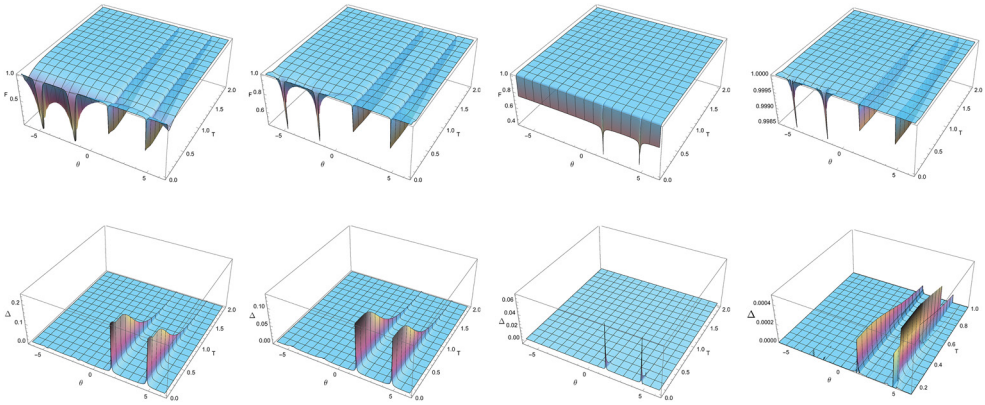


Figure C1. Fidelity (top) and $\Delta(\rho, \rho')$ (bottom) for the many-body states $\varrho^{(0)}$ and $\varrho^{(1)}$, and the single-particle states $\rho^{(0)}$ and $\rho^{(1)}$, for the AIII symmetry class. $\delta\theta = \theta' - \theta = 0.01$ and $\delta T = T' - T = 0.01$. In the case of the state $\rho^{(1)}$, the quantity Δ is highly oscillating for temperatures close to 0 in the neighbourhood of the critical points, therefore we show the results for a range of temperatures where these numerical instabilities are less prominent.

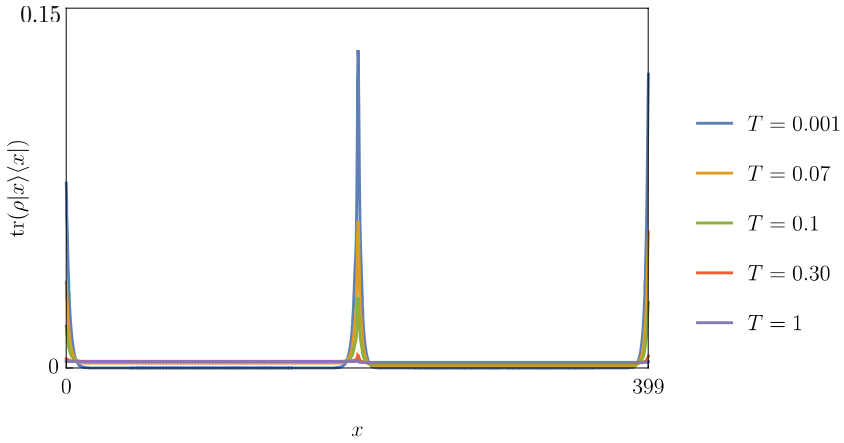
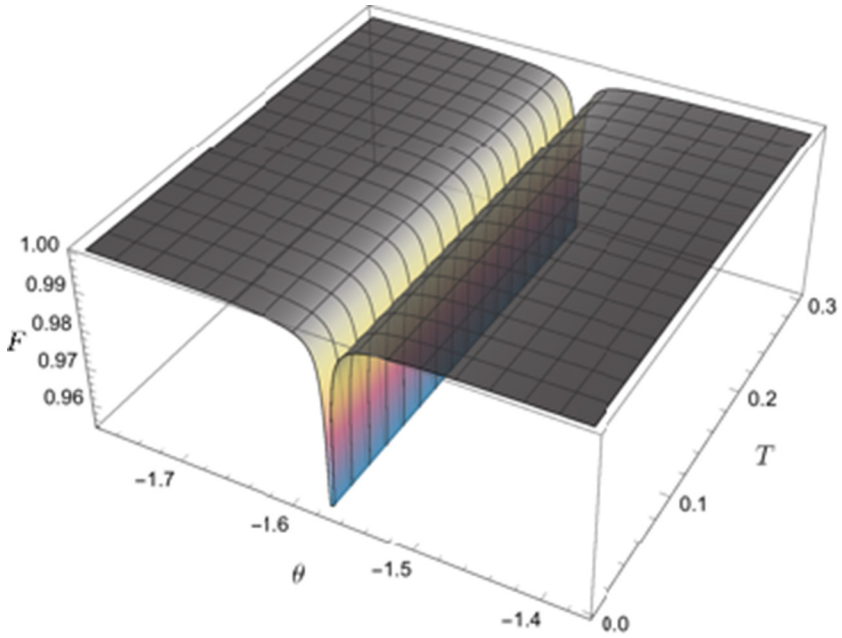


Figure C2. Position probability distribution of the QW as a function of the sites, $\mathrm{tr}(\mathrm{e}^{-H/T}|x\rangle\langle x|)/Z$. The Hamiltonian H is obtained by varying θ along x through a step-like function [28]. The domain wall is centred in the middle of the line. Periodic boundary conditions are taken, hence the edge state at the boundary.

The behaviour of the edge states, as temperature increases, is shown in figure C2.

Appendix D. Comment on figure 1

In the fidelity plot for the case of many-body $\varrho^{(1)}$ state, we observe a step-like behaviour at the critical point $(\theta_c, k_c) = (-\frac{\pi}{2}, \pi)$. This is due to numerical instabilities occurring close to zero temperature. In the following figure, we plot with higher precision ($\delta\theta = \delta T = \delta k = 0.001$, instead of 0.01) a small region $T \in [0.003, 0.3]$ and $\theta \in [-\frac{\pi}{2} - 0.2, -\frac{\pi}{2} + 0.2]$ around this 'step'. We observe that this behaviour is not present anymore, confirming that it is indeed a



numerical instability and it is not associated to a physical phenomenon. Since the higher precision is relevant only around this single point (θ_c, k_c), and the higher precision calculations are time consuming (it took about two days to obtain this small fraction of the whole plot), we kept the lower precision plots in figure 1.

References

[1] Haldane F D M 1988 Model for a quantum hall effect without Landau levels: Condensed-matter realization of the 'parity anomaly' *Phys. Rev. Lett.* **61** 2015–8

[2] Thouless D J, Kohmoto M, Nightingale M P and den Nijs M 1982 Quantized Hall conductance in a two-dimensional periodic potential *Phys. Rev. Lett.* **49** 405–8

[3] Berry M V 1984 Quantal phase factors accompanying adiabatic changes *Proc. R. Soc.* A **392** 45–57

[4] Zeng B, Chen X, Zhou D-L and Wen X-G 2015 Quantum information meets quantum matter-from quantum entanglement to topological phase in many-body systems (arXiv:1508.02595)

[5] Ando Y 2013 Topological insulator materials *J. Phys. Soc. Japan* **82** 102001

[6] Wen X G 1991 Gapless boundary excitations in the quantum hall states and in the chiral spin states *Phys. Rev.* B **43** 11025–36

[7] Ryu S and Hatsugai Y 2002 Topological origin of zero-energy edge states in particle-hole symmetric systems *Phys. Rev. Lett.* **89** 077002

[8] Uhlmann A 1989 On berry phases along mixtures of states *Ann. Phys., Lpz.* **501** 63–9

[9] Andersson O, Bengtsson I, Ericsson M and Sjöqvist E 2016 Geometric phases for mixed states of the kitaev chain *Phil. Trans. R. Soc.* A **374** 20150231

[10] Viyuela O, Rivas A and Martin-Delgado M A 2014 Uhlmann phase as a topological measure for one-dimensional fermion systems *Phys. Rev. Lett.* **112** 130401

[11] Viyuela O, Rivas A and Martin-Delgado M A 2014 Two-dimensional density-matrix topological fermionic phases: topological Uhlmann numbers *Phys. Rev. Lett.* **113** 076408

[12] Viyuela O, Rivas A and Martin-Delgado M A 2015 Symmetry-protected topological phases at finite temperature *2D Mater.* **2** 034006

[13] Huang Z and Arovas D P 2014 Topological indices for open and thermal systems via Uhlmann's phase *Phys. Rev. Lett.* **113** 076407

[14] Rivas A, Viyuela O and Martin-Delgado M A 2013 Density-matrix chern insulators: finite-temperature generalization of topological insulators *Phys. Rev.* B **88** 155141
[15] Hamma A, Ionicioiu R and Zanardi P 2005 Ground state entanglement and geometric entropy in the Kitaev model *Phys. Lett.* A **337** 22–8
[16] Hamma A, Zhang W, Haas S and Lidar D A 2008 Entanglement, fidelity, and topological entropy in a quantum phase transition to topological order *Phys. Rev.* B **77** 155111
[17] Halász G B and Hamma A 2012 Probing topological order with Rényi entropy *Phys. Rev.* A **86** 062330
[18] Abasto D F, Hamma A and Zanardi P 2008 Fidelity analysis of topological quantum phase transitions *Phys. Rev.* A **78** 010301
[19] Zanardi P and Paunković N 2006 Ground state overlap and quantum phase transitions *Phys. Rev.* E **74** 031123
[20] Zhao J-H and Zhou H-Q 2009 Singularities in ground-state fidelity and quantum phase transitions for the Kitaev model *Phys. Rev.* B **80** 014403
[21] Oliveira T P and Sacramento P D 2014 Entanglement modes and topological phase transitions in superconductors *Phys. Rev.* B **89** 094512
[22] Paunković N, Sacramento P D, Nogueira P, Vieira V R and Dugaev V K 2008 Fidelity between partial states as a signature of quantum phase transitions *Phys. Rev.* A **77** 052302
[23] Gu S-J, Kwok H-M, Ning W-Q and Lin H-Q 2008 Fidelity susceptibility, scaling, and universality in quantum critical phenomena *Phys. Rev.* B **77** 245109
[24] Maz D and Hamma A 2012 Topological order, entanglement, and quantum memory at finite temperature *Ann. Phys., NY* **327** 2096–111
[25] Zanardi P, Giorda P and Cozzini M 2007 Information-theoretic differential geometry of quantum phase transitions *Phys. Rev. Lett.* **99** 100603
[26] Paunković N and Vieira V R 2008 Macroscopic distinguishability between quantum states defining different phases of matter: fidelity and the Uhlmann geometric phase *Phys. Rev.* E **77** 011129
[27] Mera B, Vlachou C, Paunković N and Vieira V R 2017 The Uhlmann connection in fermionic systems undergoing phase transitions *Phys. Rev. Lett.* **119** 015702
[28] Kitagawa T, Rudner M S, Berg E and Demler E 2010 Exploring topological phases with quantum walks *Phys. Rev.* A **82** 033429
[29] Schnyder A P, Ryu S, Furusaki A and Ludwig A W W 2008 Classification of topological insulators and superconductors in three spatial dimensions *Phys. Rev.* B **78** 195125
[30] Kitaev A 2009 Periodic table for topological insulators and superconductors *AIP Conf. Proc.* **1134** 22
[31] Kitagawa T 2012 Topological phenomena in quantum walks: elementary introduction to the physics of topological phases *Quantum Inf. Process.* **11** 1107–48
[32] Kitagawa T, Broome M A, Fedrizzi A, Rudner M S, Berg E, Kassal I, Aspuru-Guzik A, Demler E and White A G 2012 Observation of topologically protected bound states in photonic quantum walks *Nat. Commun.* **3** 882
[33] Rakovszky T, Asbóth J K and Alberti A 2016 Detecting topological invariants in chiral symmetric insulators via losses *Phys. Rev.* B **95** 201407
[34] Groh T, Brakhane S, Alt W, Meschede D, Asbóth J K and Alberti A 2016 Robustness of topologically protected edge states in quantum walk experiments with neutral atoms *Phys. Rev.* A **94** 013620
[35] Mugel S, Celi A, Massignan P, Asbóth J K, Lewenstein M and Lobo C 2016 Topological bound states of a quantum walk with cold atoms *Phys. Rev.* A **94** 023631
[36] Kitagawa T, Berg E, Rudner M and Demler E 2010 Topological characterization of periodically driven quantum systems *Phys. Rev.* B **82** 235114
[37] Asbóth J K 2012 Symmetries, topological phases, and bound states in the one-dimensional quantum walk *Phys. Rev.* B **86** 195414
[38] Asbóth J K and Obuse H 2013 Bulk-boundary correspondence for chiral symmetric quantum walks *Phys. Rev.* B **88** 121406
[39] Shirai T, Mori T and Miyashita S 2015 Condition for emergence of the Floquet–Gibbs state in periodically driven open systems *Phys. Rev.* E **91** 030101
[40] Shirai T, Thingna J, Mori T, Denisov S, Hnggi P and Miyashita S 2016 Effective Floquet–Gibbs states for dissipative quantum systems *New J. Phys.* **18** 053008
[41] Barkhofen S, Nitsche T, Elster F, Lorz L, Gabris A, Jex I and Silberhorn C 2016 Measuring topological invariants and protected bound states in disordered discrete time quantum walks (arXiv:1606.00299v1)

[42] Vieira V R 2010 Quantum information and phase transitions: fidelity and state distinguishability *J. Phys. Conf. Ser.* **213** 012005
[43] Rigol M, Dunjko V, Yurovsky V and Olshanii M 2007 Relaxation in a completely integrable many-body quantum system: an *ab initio* study of the dynamics of the highly excited states of 1D lattice hard-core bosons *Phys. Rev. Lett.* **98** 050405
[44] Rigol M, Muramatsu A and Olshanii M 2006 Hard-core bosons on optical superlattices: dynamics and relaxation in the superfluid and insulating regimes *Phys. Rev.* A **74** 053616
[45] Vidmar L and Rigol M 2016 Generalized gibbs ensemble in integrable lattice models *J. Stat. Mech.* 064007

Journal of Physics A:
Mathematical and
Theoretical

PAPER

Oblivious transfer based on single-qubit rotations

To cite this article: João Rodrigues *et al* 2017 *J. Phys. A: Math. Theor.* **50** 205301

View the article online for updates and enhancements.

You may also like

- Encrypted machine learning of molecular quantum properties
Jan Weinreich, Guido Falk von Rudorff and O Anatole von Lilienfeld
- Efficient one-out-of-two quantum oblivious transfer based on four-coherent-state postselection protocol
I-Chao Chen, Tzonelih Hwang and Chuan-Ming Li
- Research on the security Oblivious Transfer protocol based on ECDDH
Xiaotian Wang and Zichen Li

This content was downloaded from IP address 147.91.81.84 on 19/09/2025 at 14:47

IOP Publishing Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **50** (2017) 205301 (25pp) https://doi.org/10.1088/1751-8121/aa6a69

Oblivious transfer based on single-qubit rotations

João Rodrigues[1,2], Paulo Mateus[1,2], Nikola Paunković[1,2] and André Souto[3,4]

[1] Instituto de Telecomunicações, Av. Rovisco Pais 1049-001, Lisboa, Portugal
[2] Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1049-001, Lisboa, Portugal
[3] LaSIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal Campo Grande, 1749-016 Lisboa, Portugal
[4] Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

E-mail: ansouto@fc.ul.pt

Received 8 July 2016, revised 2 March 2017
Accepted for publication 31 March 2017
Published 21 April 2017



Abstract

We present a bit-string quantum oblivious transfer protocol based on single-qubit rotations. Our protocol is built upon a previously proposed quantum public-key protocol and its practical security relies on the laws of Quantum Mechanics. Practical security is reflected in the fact that, due to technological limitations, the receiver (Bob) of the transferred bit-string is restricted to performing only 'few-qubit' coherent measurements. We also present a single-bit oblivious transfer based on the proposed bit-string protocol. The protocol can be implemented with current technology based on optics.

Keywords: oblivious transfer, quantum cryptography, secure multiparty computation

(Some figures may appear in colour only in the online journal)

1. Introduction

Since the success of quantum cryptography, which allowed for the exchange of secret keys [1–3], and whose security is based on the laws of physics, instead of the unproven mathematical conjectures, a hope appeared for designing various quantum protocols with improved security, with respect to their classical counterparts. One of the basic protocols used in building complex multiparty security schemes is the *oblivious transfer (OT) protocol* (often referred to as all-or-nothing OT).

 © 2017 IOP Publishing Ltd Printed in the UK

OT can be seen as a game played by two parties, Alice and Bob. Alice has many secrets that wishes to share with Bob in such a way that at the end, on average, Bob learns half of those secrets and Alice does not know which secrets Bob really knows. Each instance of this protocol, used to reveal in half of the cases Alice's secret, is the oblivious transfer protocol.

OT consists of two distinct phases: (i) the transferring phase, during which Alice sends an encoded secret information to Bob; (ii) the opening phase, during which Alice reveals enough information so that Bob can decode the secret with probability 1/2. Note that Bob knows if he got the message or not.

OT is said to be secure if the following properties hold: (i) the protocol is *concealing*, i.e. before the opening phase, Bob is not able to learn the message sent by Alice, while after the opening phase Bob learns the message with probability 1/2; (ii) the protocol is *oblivious*, i.e. after the opening phase, Alice remains oblivious to whether or not Bob got the message.

Rabin was the first to formally present an oblivious transfer protocol in 1981 [4]. The security of Rabin's OT relies on the fact that factoring large integers is not known to be possible to perform in polynomial time on classical computers. Later, Even, Goldreich and Lempel presented a variation of this scheme called 1-*out-of*-2 *oblivious transfer* [5]. The difference to Rabin's OT is that Alice sends two messages and Bob gets only one of the two with equal probability (again, Alice does not know which message Bob decoded). Although differently defined, Crépeau showed that when the messages are single bits the two flavours of oblivious transfer protocols are equivalent, in the sense that one can be built out of the other and vice versa [6]. Furthermore, one can build an 1-out-of-2 oblivious transfer protocol that transmits bit-string messages from 1-out-of-2 oblivious transfer protocol for single bits [7–9].

The oblivious transfer is a building block of more complex security protocols [7, 10, 11] using Yao's garbled circuits [12], and various secure multiparty computation schemes [13–15].

Another cryptographic primitive used in designing more complex secure protocols is bit commitment [16]. Although it is not possible to construct an OT protocol out of a bit commitment [17] it was shown that bit commitment can be reduced to 1-out-of-2 bit oblivious transfer protocol [18]. In figure 1 we schematically present the classical reductions between the above discussed cryptographic primitives.

Due to the advance of quantum computation and quantum information, the development of cryptographic applications resilient to quantum adversaries has been extensively studied in the last decades. Wiesner launched the field of quantum cryptography in 1969 by presenting notions such as quantum money and quantum multiplexing (and only managed to publish his results in 1983 [19]), the latter being essentially a quantum counterpart of a 1-out-of-2 oblivious transfer protocol.

Further developing Wiesner ideas, Bennett and Brassard presented the well-known BB84 quantum key distribution protocol [1], which was subsequently showed to be unconditionally secure [20–23], while its classical counterparts are only computationally secure. By 'unconditionally secure' it is meant 'secure, provided the agents have unlimited power allowed by the laws of physics'. Nevertheless, inevitable imperfections in implementations lead to various practical levels of security. In the current study, we analyse the theoretical security of our protocol, while leaving the details of particular implementations for future work. Another example of a protocol whose quantum realisation outperforms its classical counterpart is the recently proposed contract signing protocol [24].

Despite these positive results, a number of no-go theorems imposed limits to quantum cryptography. Independently, Lo and Chau [25], and Mayers [26], showed that unconditionally secure quantum bit commitment protocol is impossible, within the scope of non-relativistic physics (for unconditionally secure protocols that use relativistic effects, see [27–29]). Subsequently, Lo [30] proved similar no-go theorem for all 'one-sided two-party

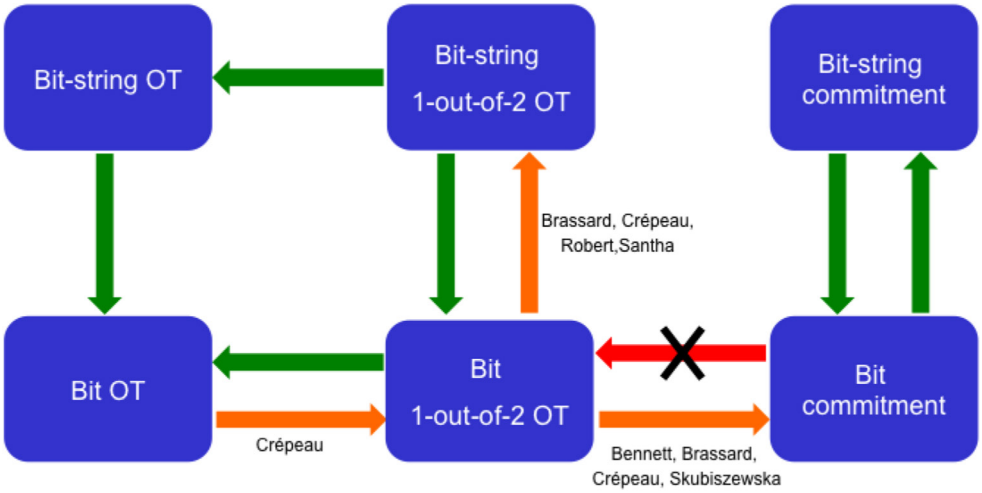


Figure 1. Classical reductions between cryptographic primitives. The green arrows represent straightforward reductions; the orange ones are non-trivial reductions; the red one is the impossible implication.

computations' protocols. An immediate consequence of this result is the impossibility of having secure 1-out-of-2 oblivious transfer. The alternative, ensuring practical security of such protocols, is to consider noisy or bounded memories [29, 31–37]. Recently, a (quantum) computationally secure version of oblivious transfer protocol was presented in [38].

Following the classical equivalence [6] between the two flavours of oblivious transfer, one might conclude that the impossibility of having unconditionally secure 1-out-of-2 oblivious transfer would imply the same for oblivious transfer. But the rules of quantum physics present a wider range of possibilities, thus potentially compromising classical reduction schemes. Namely, as to build the 1-out-of-2 oblivious transfer one has to run several oblivious transfer protocols as black boxes, the possibility of the so-called coherent attacks—joint quantum measurements on several black boxes—arises. Thus, having secure quantum oblivious transfer protocol does not necessarily mean that it is possible to construct secure 1-out-of-2 oblivious transfer. Indeed, He and Wang recently showed that in quantum domain the various types of oblivious transfer are no longer equivalent [39] and constructed a secure quantum single-bit oblivious transfer [40] using entanglement. Consequently, classical reductions of single-bit to a bit-string protocols are also compromised in the quantum setting and need to be re-examined. Recent example of constructing a secure quantum bit-string commitment protocol [41], despite the above mentioned no-go theorems for single-bit commitment [25, 26] is yet another example of invalidity of classical reductions (see also a quantum bit-string generation protocol [42]). Therefore, a need of explicitly constructing quantum bit-string oblivious transfer protocol which is not based on classical reductions mentioned above arises [6–9]. In figure 2 we present quantum reductions between cryptographic primitives. Note that, despite the mentioned controversial (and not widely accepted within the community) results by He and Wang [39, 40], the recent construction of a single-bit 1-out-of-2 oblivious transfer out of an ordinary (i.e. all-or-nothing) OT [43] leaves the question of the equivalence between the two flavours of the protocol open.

In this paper, we present a quantum oblivious transfer protocol for bit-strings, based on the recently proposed public-key cryptosystem [44]. Each bit of the string to be transferred is encoded in a quantum state of a qubit, in such a way that states corresponding to bit values

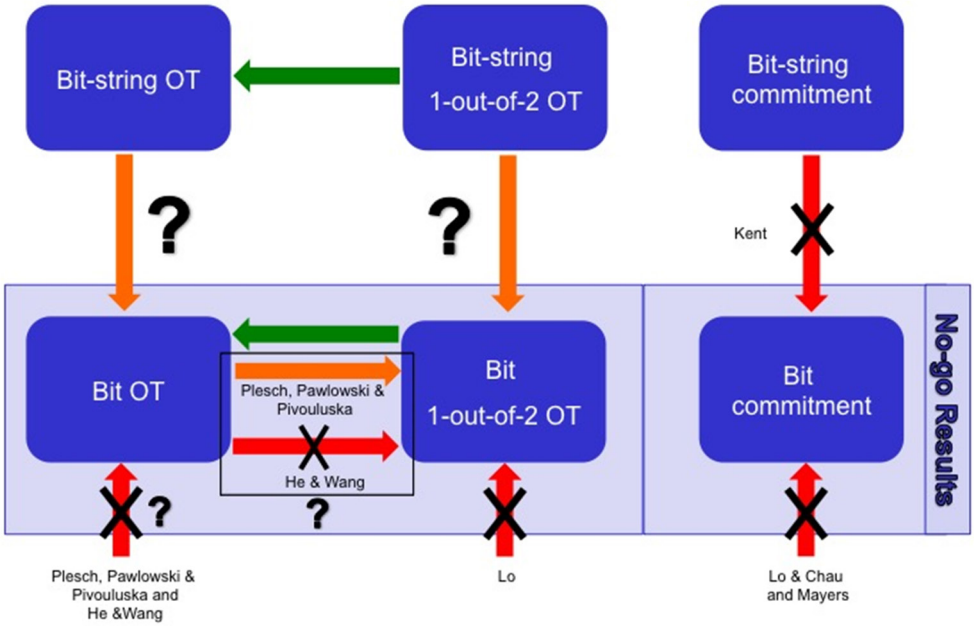


Figure 2. Quantum reductions between cryptographic primitives and no-go theorems. The green arrows represent straightforward reductions; the orange ones are non-trivial reductions; the red ones are the impossible implications. Open questions are represented with a question marks.

0 and 1 form an orthonormal basis. The key point of the protocol is that for each qubit, the encoding basis is chosen at random, from some discrete set of bases.

In the OT protocol, the two agents do not trust each other, and so no authentication is required at this level. OT is a protocol that implements a channel to be used as a building block of more complex multi-party computations. Nevertheless, at this higher level, classical authentication might be required, which can be achieved by exchanging signed messages (using pre-shared keys) via OT (see for example [45], section 2.3.2.3, p 48). Alternatively, OT can be used to achieve the very authentication between the agents [46]. In appendix A, we present details of the authentication mechanism performed at the level of multi-party computation.

The paper is organised as follows. In the next section, we present our OT protocol, while in section 3 we analyse the protocol's security. Finally, in section 4 we summarise the results and present possible future lines of research. In appendix B we present the technical details (notation, definitions and results) used in the paper.

2. The protocol

In this section, we present the protocol that achieves oblivious transfer of a bit-string message from Alice to Bob. The scheme uses hash functions which allow to certify if after the opening phase Bob got the message or not. A hash function produces a *digest* of a message—a string of smaller size—such that: (i) the probability of generating at random strings with the same hash value is negligible; (ii) the hash values are almost uniformly distributed over the set of all possible digests.

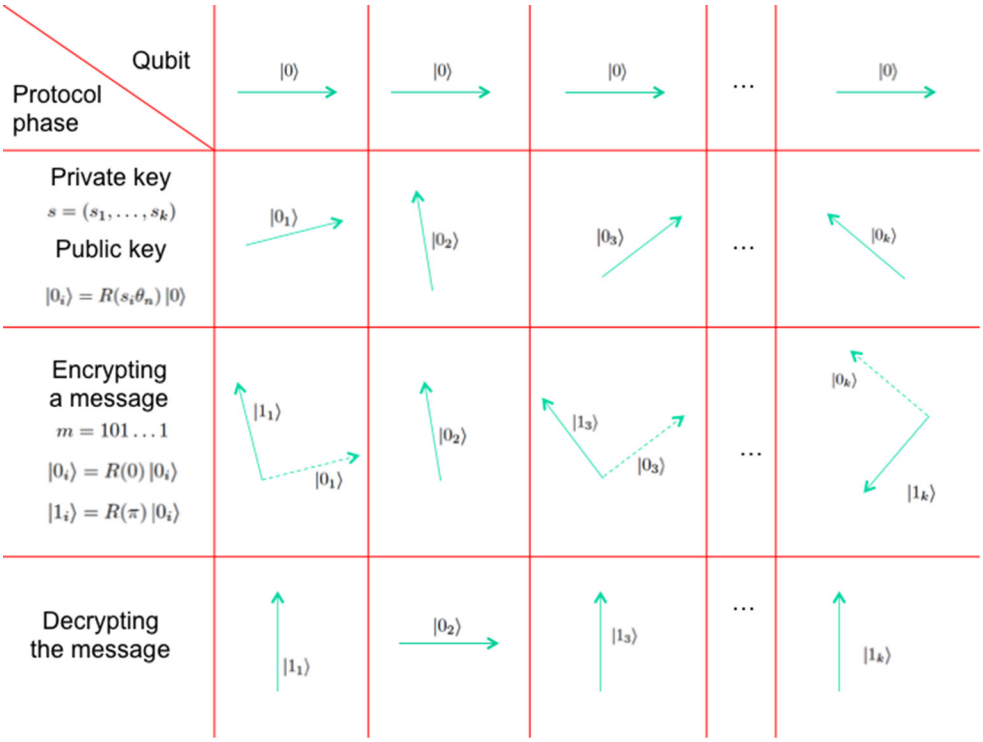


Figure 3. Schematic description of the transferring phase of our oblivious transfer protocol for messages of length k. The full arrows represent the actual states of qubits, while the dashed arrows in the last two lines (encryption of a message) represent $|0_i\rangle$ states.

Our protocol is based on the public-key cryptosystem [44], and can be briefly summarised as follows (for more details, see appendix B). Given a reference, so-called computational basis $\beta_0 = \{|0\rangle, |1\rangle\}$, Alice first encodes each bit m_i of the message (a string of bits) $\mathbf{m} = m_1 \ldots m_k$ into the state $|m_i\rangle$ of the corresponding qubit. Then, she randomly chooses a bit value a, and for each m_i a rotation angle φ_i (taken from a given set of angles Φ), and rotates $|m_i\rangle$ by $(-1)^a\varphi_i$. Finalising the transferring phase, she sends the qubits to Bob. Note that for each qubit i the encoding quantum states

$$|0_i^{(a)}\rangle = R((-1)^a\varphi_i)|0\rangle \tag{1}$$

$$|1_i^{(a)}\rangle = R((-1)^a\varphi_i)|1\rangle = R(\pi)|0_i^{(a)}\rangle, \tag{2}$$

where rotations $R(\varphi)$ are defined by $R(\varphi)|0\rangle = \cos(\varphi/2)|0\rangle + \mathrm{i}\sin(\varphi/2)|1\rangle$, are mutually orthogonal and hence fully distinguishable, provided one knows the direction a and the angle φ_i of the rotation. Therefore, Bob cannot decipher the message $\mathbf{m}$, unless given additional information about the encoding bases $\beta_i = \{|0_i^{(a)}\rangle, |1_i^{(a)}\rangle\}$. In figure 3, we present a schematic description when the length of the message to be transferred is k.

In the opening phase, Alice provides Bob with such (partial) information: she sends the so-called secret key, a string $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_k)$ of rotation angles, but *not* the rotation direction a. Oblivious to the rotation direction, Bob can only guess it, which he will get correctly in 50% of the cases.

Encrypted in quantum states of qubits, Alice sends the message $\mathbf{m}$, together with its digest $\mathbf{d} = h(\mathbf{m})$, given by a suitably chosen hash function $h : \{0,1\}^k \to \{0,1\}^\omega$, with $\omega = \lfloor\sqrt{k}\rfloor$. Upon decrypting the states of qubits sent by Alice, Bob recovers a string which is a concatenation of the form $\mathbf{m}'\mathbf{d}'$. Note that $\mathbf{m}'$ and $\mathbf{d}'$ are not necessarily the message $\mathbf{m}$ and its hash value $\mathbf{d} = h(\mathbf{m})$. Bob checks if $\mathbf{d}' = h(\mathbf{m}')$. If so, he is convinced that the received message $\mathbf{m}'$ is indeed the intended message $\mathbf{m}$ (for technical details, see section 3).

Below, we present a rigorous description of our bit-string OT protocol. The parameters that determine the protocol's security levels are $k, n \in \mathbb{N}$, with k being the length of the message, while n defines the 'elementary rotation angle' $\theta_n = \pi/2^{n-1}$. The rotation angles are then given by $\varphi_i = s_i\theta_n$, where $s_i \in \{0,\ldots,2^n-1\}$. For the reasons of simplicity, we will use the term 'secret key' for both the string of rotation angles $\boldsymbol{\varphi} = (\varphi_1,\ldots,\varphi_{k+\omega})$, as well as the string of numbers $\mathbf{s} = (s_1,\ldots,s_{k+\omega})$ that determine them. Before the opening phase, all the information about the protocol parameters are private to Alice, while the only public information is the hash function h (note that knowing the hash function, one knows k and ω as well). At the beginning of the opening phase, Alice (publicly) announces to Bob n and $\mathbf{s}$.

Protocol 1 (Bit-string OT).

Security parameters: $k, n \in \mathbb{N}$, *with* $\theta_n = \pi/2^{n-1}$;
Message to transfer: $\mathbf{m} = m_1 \ldots m_k$;
Hash function: $h : \{0,1\}^k \to \{0,1\}^\omega$, *with* $\omega = \lfloor\sqrt{k}\rfloor$;
Secret key: $\mathbf{s} = (s_1,\ldots,s_{k+\omega})$, *where each* $s_i \in \{0,\ldots,2^n-1\}$.

Transferring phase:

(i) *Alice chooses uniformly at random the hash function h and a bit* $a \in \{0,1\}$ *and prepares the following state (with* $\mathbf{d} = h(\mathbf{m})$*, and* d_i *being its ith bit):*

$$|\psi\rangle = \bigotimes_{i=1}^{k} R(m_i\pi + (-1)^a \times s_i\theta_n)|0\rangle \bigotimes_{i=1}^{\omega} R(d_i)\pi + (-1)^a \times s_{i+k}\theta_n|0\rangle \tag{3}$$

$$= \left(\bigotimes_{i=1}^{k}\left[\cos\left(\frac{m_i\pi + (-1)^a \times s_i\theta_n}{2}\right)|0\rangle + \sin\left(\frac{m_i\pi + (-1)^a \times s_i\theta_n}{2}\right)|1\rangle\right]\right) \otimes \left(\bigotimes_{i=1}^{\omega}\left[\cos\left(\frac{d_i\pi + (-1)^a \times s_{i+k}\theta_n}{2}\right)|0\rangle + \sin\left(\frac{h_i\pi + (-1)^a \times s_{i+k}\theta_n}{2}\right)|1\rangle\right]\right) \tag{4}$$

(ii) *Alice sends the state* $|\psi\rangle$ *to Bob.*

Opening phase:

(iii) *Alice sends to Bob the secret key* $\mathbf{s} = (s_1,\ldots,s_{k+\omega})$*, the security parameter n.*
(iv) *Bob checks if* $\mathbf{s}$ *is likely to be a possible output of a random process.*
(v) *Bob chooses uniformly at random* $a' \in \{0,1\}$ *and applies* $R((-1)^{a'}s_i\theta_n)$ *to each qubit of* $|\psi\rangle$*.*
(vi) *Bob applies the measurement operator* $M^{\otimes(k+\omega)} = (0 \times |0\rangle\langle 0| + 1 \times 1|1\rangle\langle 1|)^{\otimes(k+\omega)}$.
(vii) *Let* $\mathbf{m}'\mathbf{d}'$ *be the message that Bob recovers (notice that here* $\mathbf{d}'$ *is a bit-string, a potential value of the hash, and not a function itself). He checks if* $\mathbf{d}' = h(\mathbf{m}')$*. If that is the case then Bob is almost sure that* $\mathbf{m}' = \mathbf{m}$*, otherwise he knows that* $\mathbf{m}'$ *is not the correct message.*

Notice that knowing $h(\mathbf{m})$ can potentially reveal the whole set $A_{\mathbf{m}}$ of the strings mapped to the same value of hash. Knowing $A_{\mathbf{m}}$ decreases Bob's uncertainty about the unknown string $\mathbf{m}$, thus effectively revealing $\omega = \lfloor\sqrt{k}\rfloor$ bits of information about string $\mathbf{m}$. This information may help Bob to increase the probability of finding $\mathbf{m}$, thus compromising the security of the protocol. Therefore we encrypt both the message $\mathbf{m}$ and $h(\mathbf{m})$ into a quantum state sent by Alice. Since, in order to confirm that he obtained the message $\mathbf{m}$, Bob needs to learn the value $h(\mathbf{m})$ as well, one can consider the concatenated string $\mathbf{m}h(\mathbf{m})$ as a message to be transferred. For simplicity, in the rest of the paper we will denote $\mathbf{m}h(\mathbf{m})$ as a single message $\mathbf{m}$ to be transferred. Note though, that there are correlations between the message $\mathbf{m}$ and the value $h(\mathbf{m})$, which might become relevant for the Concealing property, and in particular for achieving the Probabilistic transfer, after Bob learns the particular function $\mathbf{m}$ chosen by Alice. We will address this issue when discussing the above mentioned cases.

In Step 4 Bob checks if the secret key s was indeed randomly chosen. By encoding s_i's into binary numbers Alice has to provide an $n \times (k + \omega)$ long bit-string produced by a fair coin. A number of possible tests of random-number generators exist in literature, such as χ^2, Kolmogorov–Smirnov, serial correlation, Two-level, K-distributivity, serial and spectral tests (for more details, see [47], chapter 27). Step 4 of the protocol is used to overcome the hypothetical chance of Alice to cheat by sending particular elements s_i of the secret key s which allow Bob to recover the message with probability close to 1. Notice that if s_i is such that the angle of rotation is $\varphi_i = s_i\theta_n/2 = 0$, or $\varphi_i = s_i\theta_n/2 = \pi/2$, then Bob will with certainty get the correct bit value m_i. Therefore, if the elements s_i of the secret key were close to 0 or $\pi/2$, Alice would know with probability significantly higher than 1/2 that Bob received the message $\mathbf{m}$ (for a detailed analysis of possible cheating strategies of Alice, see below the proof of the obliviousness criterion). If all s_i's were indeed chosen uniformly at random, than significant portion of them would not be close to 0 nor $\pi/2$, preventing Alice from cheating.

Nevertheless, for the protocol to be secure, a much simpler criterion can be used, one that is satisfied whenever a string is indeed produced uniformly at random. If Alice chooses each s_i uniformly at random, then on average half of such choices satisfy $\varphi_i = s_i\theta_n/2 \in [\pi/8, 3\pi/8] \cup [5\pi/8, 7\pi/8]$. These s_i's are already far enough from 0 and $\pi/2$ to secure the protocol against cheating Alice. For a detailed discussion on the degree of Bob's confidence against cheating strategies of Alice, see below the proof of the obliviousness criterion.

Typically, bit-string protocols achieve required security levels for *suitably large* lengths of messages, as is the case in Kent's protocol [41], for example (for more details see section 3). Nevertheless, having a secure bit-string OT protocol by encrypting single bits by sufficiently large messages, one can design secure single-bit OT. Below, we present a simple way of using our protocol to achieve oblivious transfer of a single bit b by sending a bit-string message $\mathbf{m}$.

Protocol 2 (Single-bit oblivious transfer).

Message to transfer: b;
Security parameters: k *and* n;

(i) *Alice chooses bit b.*
(ii) *Alice chooses a 2k-bit message* $\mathbf{m}$*, such that* $\bigoplus_{i=1}^{k} m_i \times m_{k+i} = b$.
(iii) *Alice and Bob perform protocol 1.*
(iv) *If Bob had got the right message* $\mathbf{m}$*, then he performs* $\bigoplus_{i=1}^{k} m_i \times m_{k+i} = b$*. Otherwise, he cannot recover the bit.*

3. Security analysis

In this section, we analyse the security of our oblivious transfer protocol. Oblivious transfer has to satisfy the following four properties (the first express the correctness while the last three assure the security of the protocol):

Soundness: If both Alice and Bob are honest, then with probability 1/2 Bob will obtain the right message. While Bob knows if he got the right message or not, Alice is oblivious of that fact.
Concealingness: If Alice is honest Bob cannot learn the whole message that Alice meant to send before the opening phase (the protocol is concealing).
Probabilistic transfer: After the opening phase, Bob cannot learn the whole message in more than 50% of the cases (with probability higher than 1/2).
Obliviousness: If Bob is honest then Alice does not know if Bob received the message—she can only guess with probability 1/2 (the protocol is oblivious).

In case of bit-string protocols, the probability 1/2 that appears in the above definition of Soundness and Probabilistic transfer properties (but not Obliviousness) is relaxed to $1/2 + \varepsilon(k)$, where k is the length of the message and $\varepsilon : \mathbb{N} \to \mathbb{R}$ is a negligible function, i.e. for every positive polynomial p there exists a $k_0 \in \mathbb{N}$ such that for all $k > k_0$, $\varepsilon(k) \leqslant 1/p(k)$ (for the definition and a detailed overview of the use of the notion of negligible functions in cryptography, see for example [48]).

Note that Bob can get *a part* of the message $\mathbf{m}$ with certainty, as long as the number of additional bits that he has to guess is larger than a polylogarithmic function of k (a function is polylogarithmic if it can be written in the form $f(x) = \sum_i c_i \log^{p_i}(x)$, for fixed constants c_i and powers p_i, see for example [49]). This means that the uncertainty of the entire message is greater than any polylogarithmic function of k. For instance, in the case a cheating Bob can learn all but $\sqrt{k}$ bits of the message, the probability of him guessing properly the *whole* message is $2^{-\sqrt{k}}$ (i.e. Bob's uncertainty of the whole message is $\sqrt{k}$). This way, Alice can use our protocol to send $\sqrt{k}$ bits of information of a shorter string $\mathbf{i}$ whenever the transfer is successfully achieved; otherwise, Bob's probability to learn the intended information $\mathbf{i}$ is negligible (e.g. of the order of $2^{-\sqrt{k}}$; note that the quadratic increases in the resources is common and acceptable in implementation of communication protocols). To do this, she can use an error-correcting code that can correct up to $\sqrt{k}$ errors. Such code consists of cells ('codewords') each of size $2^{\sqrt{k}}$. By randomly choosing the cell, and encoding $\mathbf{i}$ in one of its $2^{\sqrt{k}}$ elements, Bob is left without virtually any knowledge of $\mathbf{i}$ whenever the OT is unsuccessful: by being able to learn $(k - \sqrt{k})$ bits of $\mathbf{m}$ he can at most identify the cell, but not $\mathbf{i}$ (the uncertainty of the received string $\mathbf{m}$ is precisely $\sqrt{k}$). Note the importance of stating the security criterion in terms of negligible functions, i.e. in terms of asymptotic behaviour, with respect to the message length (be it k of $\mathbf{m}$ or, through the error-correcting code, $\sqrt{k}$ of $\mathbf{i}$), of the probabilities to learn (parts of) the information sent, which is a standard approach in cryptography and computer science (see for example [48], where negligible functions are introduced already in the Introduction).

In general, both quantum and classical cryptographic security protocols for exchanging messages depend on several parameters, one of them being the length of the message. As a rule, such protocols are said to be secure if, provided that the other parameters are suitably chosen, the cheating probability is negligible with respect to the length of the message, which is satisfied asymptotically, i.e. for *sufficiently large* values of k.

In our case, as well as in the case of the public-key scheme presented in [44] (on which our protocol is based; see also the related recent scheme based on quantum walks [50]), one such

parameter is n, and for both protocols the level of security indeed depends on the choice of n. Nevertheless, as proven in [44], with a proper choice of n, the public-key scheme is secure against eavesdropping. Consequently, with the same choice of a proper n, our protocol is sound (correct), Concealing (before the opening phase Bob cannot learn the message sent by Alice) and achieves Probabilistic transfer (on average, Bob receives half of the messages sent by Alice). On the other hand, the last security criterion (obliviousness) does not depend on the choice of n, as shown in the respective proofs presented below.

Note that in [44], in order to further reduce the probability of a successful attack, security parameter n was treated as a part of the secret key (together with s). But it was noted that the protocol would still be secure even if n were public. In a subsequent paper [51], in which the robustness of the public-key cryptosystem introduced in [44] was further analysed, n was treated as a part of a public key, i.e. the cryptosystem is secure even if (a properly chosen) n were known. Note that in both cases, according to the above definition, the protocol is secure, but with different negligible functions $\varepsilon(k)$: when n is private, the corresponding negligible function is smaller than when n is public.

Definitions of soundness, concealingness and probabilistic transfer properties have somewhat weak requirement of Bob not being able to learn the *whole* message $\mathbf{m}$ before and in (about) half of the cases after the opening phase. According to this definition, even if Bob were able to learn all but, say the last bit of the message, the protocol would still be secure. One can adopt a stronger criterion, requiring that only with negligible probability Bob can learn *with certainty* part of $\mathbf{m}$ which would allow him to *infer*, with *non-negligible probability*, the rest of the message. In other words, to require that the acquired information cannot help Bob to *improve* his chances of guessing $\mathbf{m}$, in a sense that before and after receiving quantum (and classical) key(s), Bob's chances to correctly guess the whole $\mathbf{m}$ are negligible in the length of the message. Such security criterion is in the spirit of the one presented in the classic paper [5] introducing 1-out-of-2 oblivious transfer protocol.

Note that even in the case of the correct inference of the rest of the message, a cheating Bob can still not know *with certainty* whether the guess was correct, which is weaker position than that of an honest player who knows with certainty if the message is received. Note that by sheer guessing, Bob will always, on average, correctly infer half of the bits of $\mathbf{m}$, but without knowing which bits he guessed right. To say that Bob knows certain part of the message means to know both the values and the corresponding positions of the known bits. Thus, there is a difference between knowing the message (as required by the definition of the protocol), and being right in inferring (guessing) the message.

As we show below, regarding Concealingness, our protocol satisfies the strongest possible requirement: no part of the message can be known at all (with certainty). Indeed, upon obtaining a string of qubits from Alice, Bob can *at best* learn the whole quantum state of the system received. Since each quantum state $|\psi\rangle$ sent by an honest Alice can, for a suitable choice of the secret key $\mathbf{s} = (s_1, \ldots, s_{k+\omega})$, encode *any* string $\mathbf{md} \in \Sigma^{k+\omega}$, before the opening phase Bob is completely clueless of any part of the message sent.

Regarding Probabilistic transfer, we show that, providing that Bob is constrained to perform only 'few-qubit' coherent measurements, he can at best with probability of 1/2 learn all but the first $\sqrt{k}$ bits of the message. Nevertheless, the probability to successfully infer the missing $\sqrt{k}$ bits is $2^{-\sqrt{k}}$, i.e. negligible in the length of the message. By 'few-qubit' coherent measurements we mean that Bob cannot perform a joint measurement on more than a given finite number of qubits, a reasonable practical constraint for today's and any not-so-far future technology.

3.1. Soundness of the protocol

In the following we prove the soundness of our protocol: if both parties are honest, then with probability $1/2 + \varepsilon(k)$ Bob will get the right message, where $\varepsilon(k)$ is negligible function on the size of the message $\mathbf{m} = m_1 \ldots m_k$.

First assume that Alice and Bob had chosen to rotate the state in opposite directions, i.e. $a \neq a'$. Without loss of generality assume that Alice chooses $a = 0$, to rotate clockwise all the qubits. The qubits Alice sent to Bob are in the following state:

$$|\psi\rangle = \bigotimes_{i=1}^{k} R(m_i\pi + s_i\theta_n)|0\rangle \tag{5}$$

$$= \bigotimes_{i=1}^{k} \cos\left(\frac{m_i\pi + s_i\theta_n}{2}\right)|0\rangle + \sin\left(\frac{m_i\pi + s_i\theta_n}{2}\right)|1\rangle. \tag{6}$$

In the opening phase Bob receives from Alice the additional information, the secret key $s = (s_1, \ldots, s_k)$.

By the assumption, Bob decides to rotate each qubit received from Alice counterclockwise ($a' = 1$) by $-s_i\theta_n$. The states he gets are either $|0\rangle$ or $|1\rangle$. In fact:

$$R(-s_i\theta_n)(R(m_i\pi + s_i\theta_n)|0\rangle) = R(m_i\pi)|0\rangle \tag{7}$$

$$= \cos\left(\frac{m_i\pi}{2}\right)|0\rangle + \sin\left(\frac{m_i\pi}{2}\right)|1\rangle \tag{8}$$

$$= |m_i\rangle. \tag{9}$$

Bob measures $M = 0 \times |0\rangle\langle 0| + 1 \times |1\rangle\langle 1|$ on the above state and the result is m_i with probability 1 (note that for $m_i = 0$ we have $\cos(m_i\pi/2) = 1$, and analogously for $m_i = 1$). We conclude that if Bob chooses to rotate in the direction contrarily to Alice's choice, then with probability 1 Bob will recover the bit sent by Alice.

On the other hand, if Alice and Bob decide to rotate each qubit of the message in the same direction ($a = a'$), say clockwise, the qubits' states are transformed into ($i = 1 \ldots k$):

$$R(s_i\theta_n)(R(m_i\pi + s_i\theta_n)|0\rangle) = R(m_i\pi + 2s_i\theta_n)|0\rangle \tag{10}$$

$$= \cos\left(\frac{2s_i\theta_n + m_i\pi}{2}\right)|0\rangle + \sin\left(\frac{2s_i\theta_n + m_i\pi}{2}\right)|1\rangle \tag{11}$$

$$= |\tilde{m}_i\rangle. \tag{12}$$

If $m_i = 0$ then the above state becomes $|\tilde{m}_i\rangle = \cos(s_i\theta_n)|0\rangle + \sin(s_i\theta_n)|1\rangle$ and by measuring the qubit with $M = 0 \times |0\rangle\langle 0| + 1 \times |1\rangle\langle 1|$ Bob gets the correct answer with probability $\cos^2(s_i\theta_n)$; if $m_i = 1$ then the above state becomes $|\tilde{m}_i\rangle = -\sin(s_i\theta_n)|0\rangle + \cos(s_i\theta_n)|1\rangle$ and again Bob gets the correct bit with probability $\cos^2(s_i\theta_n)$. Hence

$$\Pr(m_i; M, |\tilde{m}_i\rangle) = \cos^2(s_i\theta_n). \tag{13}$$

Assuming that the key $\mathbf{s}$ is chosen at random, the probability of recovering the whole message by rotating in the wrong direction becomes negligible, and the expected probability of recovering message $\mathbf{m}$, when measuring $M^{\otimes k}$, on the state $|\psi'\rangle = \bigotimes_{i=1}^{k} R((-1)^{a'} s_i\theta_n)|\psi\rangle$ is:

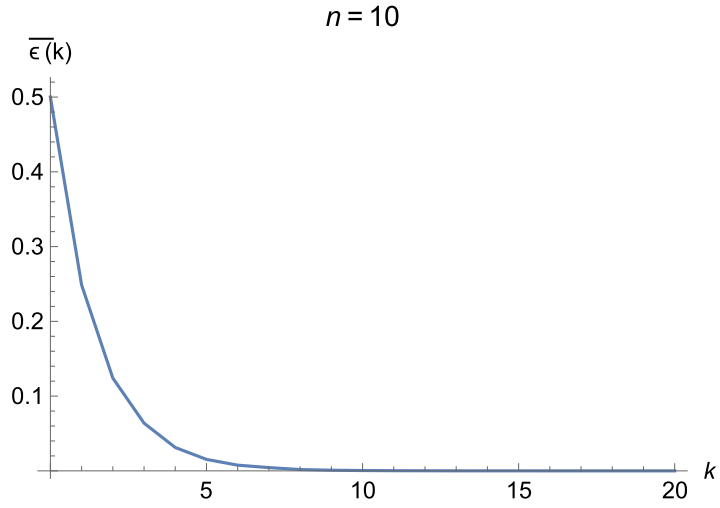


Figure 4. The expected value $\bar{\varepsilon}(k)$, for $n = 10$ and 100 randomly chosen strings $\mathbf{s}$.

$$\Pr(\mathbf{m}; M^{\otimes k}, |\psi'\rangle) = \Pr(a' \neq a) \times \Pr(\mathbf{m}|a' \neq a) + \Pr(a' = a) \times \Pr(\mathbf{m}|a' = a) \tag{14}$$

$$\leqslant \frac{1}{2} + \frac{1}{2} \prod_{i=1}^{k} \cos^2(s_i\theta_n). \tag{15}$$

The two cases, $a' \neq a$ and $a' = a$, occur both with probability 1/2. While in the first case Bob always gets $\mathbf{m}$ correctly, in the second, the probability given by the random values s_i chosen by Alice, is given by $\prod_{i=1}^{k} \cos^2(s_i\theta_n)$. To see that $\varepsilon(k) = \frac{1}{2}\prod_{i=1}^{k} \cos^2(s_i\theta_n)$ is negligible, notice that on average half of values for the rotation angles $s_i\theta_n/2$ fall in the region $[\pi/8; 3\pi/8] \cup [5\pi/8; 7\pi/8]$, for which $\cos(s_i\theta_n) < \sqrt{2+\sqrt{2}}/2$, giving the upper bound $\varepsilon(k) \leqslant 2^{-k/2}$.

Given θ_n, one can estimate the average value $\bar{\varepsilon}(k)$ as a function of the message length k, thus obtaining the security level of the protocol's soundness criterion. The current technology available at the market allows for the accuracy of the state of polarisation (SOP) of about 0.2° on Poincaré sphere (see the accuracy specifications of the 'polarization instrumentation' at [52]). This roughly corresponds to the case of $n = 10$, for which on figure 4 we plot the expectation value $\bar{\varepsilon}(k)$, for 100 randomly chosen strings $\mathbf{s}$. As one can see, already for the messages of four letters the 'access' probability $\bar{\varepsilon}(k)/2$ is of the order of 0.01.

The information received by Bob consists of two parts: one corresponding to the actual message sent by Alice, and the other corresponding to its hash value. At the end of the protocol, Bob checks if he recovered the correct message by comparing its hash value with the latter part of information received. Note that by the properties of hash functions, the probability that the hash of the first part matches the second one is negligible in the case Alice and Bob performed the same rotation (see appendix B for more detailed description of the properties of hash functions).

3.2. Concealingness of the protocol

In this section we show that if Alice is honest, the probability of Bob recovering Alice's message before the opening phase is negligible. Furthermore, after the opening phase Bob recovers the message with, up to a negligible value, probability 1/2.

The first part of the statement follows directly from the security of the public-key cryptosystem [44] (see the discussion on one-way functions and state distinguishability in appendix B), and is basically a consequence of the fact that, depending on the secret key component s_i, the same state $|\psi_i\rangle$ of a single qubit can be encrypting either a 0 or a 1: for each s_i there exists s_i' such that $|\psi_i\rangle = R(s_i\theta_n)|0\rangle$ encrypts 0, while $|\psi_i\rangle = R(s_i'\theta_n + \pi)|0\rangle$ encrypts 1. In fact, before the opening phase, our protocol is as secure as the cryptographic system underlying it.

We stress that the additional information provided by Alice, the hash function cannot help Bob recovering the message $\mathbf{m}$. In fact, below, we prove that even if Bob had access directly to the hash value $\mathbf{h}(\mathbf{m})$ this would not help him (note that since the value $\mathbf{h}(\mathbf{m})$ is encrypted makes Bob's task even harder). In the following, we provide the reasoning for a particular hash function.

Given a message $\mathbf{m}$, consider its partition into $\omega = \lfloor\sqrt{k}\rfloor$ consecutive blocks of bits $\bar{\mathbf{m}}_i$ ($i = 1, \ldots \omega$), each with length $\lfloor\sqrt{k}\rfloor$: $\mathbf{m} = \bar{\mathbf{m}}_1 \ldots \bar{\mathbf{m}}_\omega$. Each bit $d_i = h_i(\mathbf{m})$ of the hash value $\mathbf{d} = h(\mathbf{m})$ is the parity of the ith block of the message $\mathbf{m}$: $d_1 = h_1(\mathbf{m}) = m_1 \oplus \ldots \oplus m_\omega$, etc. Hence, all the bits of $\mathbf{d} = h(\mathbf{m})$ are mutually independent.

Suppose that $\mathbf{d} = h(\mathbf{m})$ allows to recover $\mathbf{m}$ with some non-negligible probability p. Then, in particular, the bit $d_1 = h_1(\mathbf{m})$ helps to recover the possible block $\bar{\mathbf{m}}_1 = m_1 \oplus \ldots \oplus m_\omega$, with the same probability p. We claim that this is impossible, assuming that the cryptographic system [44], used to design our protocol, is secure.

In fact, if a cryptographic system is secure for coding a message $\mathbf{m}$ of length k, then *a fortiori* the encryption of a polynomially shorter message, say $\bar{\mathbf{m}}_1$, is also secure. So, if $d_1 = h_1(\mathbf{m})$ would help to recover the first block with non negligible probability p then, by randomly guessing the value $d_1 = h_1(\mathbf{m})$ (that will be correct with probability 1/2), it would be possible to break the cryptosystem presented in [44] with non-negligible probability $p/2$.

One can easily describe other hash functions by considering all possible forms of dividing k elements into groups of ω elements, i.e. by using the above hash function h on the permuted message. Given a permutation $\pi \in S_k$ of length k, one can define the hash function $h_\pi(\mathbf{m}) = h_{id_k}(m_{\pi(1)} \ldots m_{\pi(k)})$, where h_{id_k} is the above h. Obviously, the concealing property is valid for all hash functions in the set $\{h_\pi | \pi \in S_k\}$.

3.3. Probabilistic transfer of the protocol

After receiving the secret key $\mathbf{s}$, Bob's description of the qubits sent by Alice is given by the mixed state (for convenience, we consider $a \in \{+, -\}$, where '$+$' stands for clockwise rotation and '$-$' otherwise):

$$\rho_B(\mathbf{s}) = \frac{1}{2} \sum_{a\in\{+,-\}} \left(\frac{1}{2}\right)^k \sum_{m_1\in\{0,1\}} \cdots \sum_{m_k\in\{0,1\}} |m_1(s_1)\rangle_a\langle m_1(s_1)| \otimes \ldots \otimes |m_k(s_k)\rangle_a\langle m_k(s_k)|, \tag{16}$$

where $|m_i(s_i)\rangle_\pm = \cos\left(\frac{m_i\pi}{2} \pm \frac{s_i\theta_n}{2}\right)|0\rangle + \sin\left(\frac{m_i\pi}{2} \pm \frac{s_i\theta_n}{2}\right)|1\rangle$. The single-qubit partial states are completely mixed, and can be written in the following suitable form: $\rho_B(s_i) = \frac{1}{2}(\rho_0(s_i) + \rho_1(s_i))$, where $\rho_{m_i}(s_i) = \frac{1}{2}(|m_i(s_i)\rangle_+\langle m_i(s_i)| + |m_i(s_i)\rangle_-\langle m_i(s_i)|)$. Note though that the overall state $\rho_B(\mathbf{s})$ is not a tensor product of single-qubit states: the rotation direction a is the same for all qubit thus correlating single-qubits. Nevertheless, if Bob is constrained to perform only few-qubit coherent measurements, these correlations, as well as the knowledge of $\mathbf{d} = h(\mathbf{m})$, cannot help him to increase the probability of learning $\mathbf{m}$.

First, we give the proof for the case of single-qubit measurements. As before, the hash function h is determined by the parity of blocks $\bar{\mathbf{m}}_i$ of size ω. Since the parity of block $\bar{\mathbf{m}}_i$ is

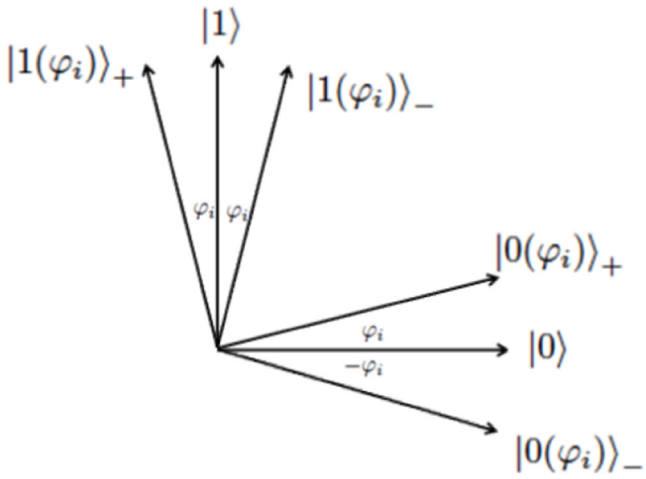


Figure 5. The optimal discrimination between the bit values 0, encoded in quantum state $\rho_0(s_i) = \frac{1}{2}(|0(s_i)\rangle_+\langle 0(s_i)| + |0(s_i)\rangle_-\langle 0(s_i)|)$, and the bit value 1, $\rho_1(s_i) = \frac{1}{2}(|1(s_i)\rangle_+\langle 1(s_i)| + |1(s_i)\rangle_-\langle 1(s_i)|)$. The optimal observable is given by the vectors from the computational basis, $|0\rangle$ for inferring the bit value 0, and $|1\rangle$ for inferring the bit value 1. Note that $\varphi_i = s_i\theta_n/2$.

completely uncorrelated to the value of each of its bits, unless we know the values of all other $\omega - 1$ remaining bits, the choice of the optimal single-qubit measurement of at least $\omega - 1$ qubits of a single block does not depend on the hash value $d_i = h_i(\mathbf{m})$.

The correlations between single-qubit states established by the same choice of the rotation direction cannot help either. A possible cheating strategy would be to, as prescribed by the protocol, randomly choose the rotation direction, and perform the corresponding measurement on first few qubits only. With probability 1/2 the choice will be right, and the bits would be correctly decrypted; with probability 1/2 though, the wrong choice would lead to wrong decryption which, in case Bob can detect it, would result in measuring the right observable on the remaining qubits. But Bob can detect the wrong choice only by comparing the results with the hash value, the parity of blocks of length ω. Thus, only upon measuring all qubits of at least one block of size ω Bob can spot the mistake. This however leaves him uncertain which, among $2^{\omega-1}$ possible messages, was the message sent by Alice, which are exponentially many on the size k of the whole message $\mathbf{m}$ (note that $\omega = \sqrt{k}$). Thus, since for each s_i the states $\rho_0(s_i)$ and $\rho_1(s_i)$ are not fully distinguishable, what Bob can do is to try to distinguish between the two states as best as possible.

The optimal probability of guessing bit's value m_i is then given by the Helstrom formula [53]:

$$\mathrm{P}_H(\rho_0(s_i), \rho_1(s_i)) = \frac{1}{2} + \frac{1}{4}\mathrm{Tr}|\rho_0(s_i) - \rho_1(s_i)| = \frac{1}{2}(1 + |\cos(s_i\theta_n)|). \qquad (17)$$

Note that the optimal observable for such measurement is the same for each possible s_i, and is given by the computational basis $\{|0\rangle, |1\rangle\}$ (see figure 5). Analogously as in the proof of soundness of the protocol, since on average half of values s_i satisfy $|\cos(s_i\theta_n)| \leqslant 1/\sqrt{2}$, we have $\varepsilon(k) \leqslant q^{-k/2}$, where $q = \frac{1}{2}(1 + 1/\sqrt{2}) < 1$. It follows that the probability to guess any non-negligible number of bits is still negligible (by 'negligible number of bits' we mean the number of bits such that, knowing their values the probability to guess the rest of $\mathbf{m}$ is negligible).

Suppose now Bob is allowed to perform at most two-qubit coherent measurements. Then, for each pair, say (s_1, s_2), the four quantum states

$$\rho_{00}(s_1 s_2) = \frac{1}{2}(\rho_{00}^{+}(s_1 s_2) + \rho_{00}^{-}(s_1 s_2)) \tag{18}$$

$$= \frac{1}{2}(|0(s_1)0(s_2)\rangle_{+}\langle 0(s_1)0(s_2)| + |0(s_1)0(s_2)\rangle_{-}\langle 0(s_1)0(s_2)|), \tag{19}$$

(and analogously for $\rho_{01}(s_1 s_2)$, $\rho_{10}(s_1 s_2)$ and $\rho_{11}(s_1 s_2)$), would also not be fully distinguishable. Therefore, the optimal strategy that Bob can adopt will produce wrong decryption, with finite error probability $q > 0$. As in the case of single-qubit measurements, this leads to negligible advantage over the 1/2 probability of recovering m, given sufficiently large k (and thus the block length $\omega = \sqrt{k}$).

Given the maximal length ℓ of the multi-qubit measurement, each block $\tilde{m}$ of length ℓ is from Bob's point of view described by the mixed state $\rho_{\tilde{\mathbf{m}}}(\tilde{\mathbf{s}}) = \frac{1}{2}(\rho_{\tilde{\mathbf{m}}}^{+}(\tilde{\mathbf{s}}) + \rho_{\tilde{\mathbf{m}}}^{-}(\tilde{\mathbf{s}})) = \frac{1}{2}(|\tilde{\mathbf{m}}(\tilde{\mathbf{s}})\rangle_{+}\langle\tilde{\mathbf{m}}(\tilde{\mathbf{s}})| + |\tilde{\mathbf{m}}(\tilde{\mathbf{s}})\rangle_{-}\langle\tilde{\mathbf{m}}(\tilde{\mathbf{s}})|)$, where $\tilde{\mathbf{s}}$ is the part of the secret key $\mathbf{s}$ corresponding to the block $\tilde{\mathbf{m}}$. As ℓ increases, the states $|\tilde{\mathbf{m}}(\tilde{\mathbf{s}})\rangle_{\pm}$ and $|\tilde{\mathbf{m}}'(\tilde{\mathbf{s}})\rangle_{\pm}$, corresponding to two different messages $\tilde{\mathbf{m}}$ and $\tilde{\mathbf{m}}'$, become increasingly distinguishable. The precise relation between the maximal length ℓ of the allowed coherent measurements and the size k of the message m will be addressed in a separate study.

To summarise, the crucial nontrivial part of the proof of the probabilistic transfer criterion restricted to few-qubit measurements is to show that the probability to learn m is negligible, i.e. it scales faster than the inverse of any polynomial on the length of the message. Since the few-qubit states of the 'maximal length' ℓ on which Bob performs the measurements are not fully distinguishable, he will infer correctly that part of the message only with probability p_ℓ strictly smaller than 1. Thus, upon repeating the finite-length few-qubit measurements $t = k/\ell$ times, the probability of correctly inferring the whole message scales exponentially to zero as $p_\ell^t = p_\ell^{k/\ell} = (\sqrt[\ell]{p_\ell})^k$, which, given the behaviour of p_ℓ (a subject of a future study) provides the explicit relation between the message length k and the maximal measurement length ℓ, for a desired security level. Thus, given the message length k and the desired security level ε (upper value for the probability to cheat), one can determine the fixed 'maximal length' ℓ of allowed multi-qubit joint measurement by solving the equation $\varepsilon = (\sqrt[\ell]{p_\ell})^k$ on ℓ. Note that by 'multi-qubit joint measurement' we mean any kind of generalised POVM allowed by the laws of quantum mechanics. Further, to perform the analysis of the relation between k and ℓ, it is enough to find a suitable upper bound of $p_\ell < 1$.

3.4. Obliviousness of the protocol

To finish the security discussion we prove that the protocol is oblivious: at the end of the protocol Alice does not know whether Bob received the right message of not.

At the end of the protocol, since Bob performs *only* local operations and measurements, *without communicating* to Alice, she has no way of knowing if Bob had chosen the right rotation, or not. Note that no entanglement *without* subsequent communication can be used to exchange/acquire information. Otherwise, one could achieve faster-than-light communication, thus explicitly violating causality and the principle of relativity. Indeed, what entanglement affects are the correlations: Bell inequalities are given in terms of various correlation functions, and the violation of local realism can be observed only upon distant observers exchange the results of their local measurements. The prominent bit-commitment cheating strategy

discovered independently by Lo and Chao [25] and Mayers [54] does involve entanglement, but in addition to performing a suitable measurement on her halves of entangled pairs, a cheating Alice has to, *subsequently*, send her measurement outcomes to Bob. Therefore, if being honest and sending the state prescribed by the Protocol, Alice cannot know if an honest Bob received the message or not.

To increase her probability of knowing if Bob received the message or not, a cheating Alice can try sending a cheating state $|\psi_{\rm ch}\rangle$ that would, with probability significantly higher than 1/2 (ideally, with probability 1), reveal $\mathbf{m}$ independently of Bob's choice of rotation (obviously, in order to maintain the 50% of Bob's success, Alice has to send a statistical mixture of different cheating states, some having the mentioned 'success probability' significantly higher than 1/2, the others significantly lower). Note that in our protocol *no entanglement* between Alice and qubits sent to Bob could possibly help her to cheat, by adapting the so-called Mayers–Lo–Chau type of attack, which indeed enables Alice to successfully cheat during the execution of a quantum bit-commitment protocol (the essence of the mentioned no-go theorem [25, 26]). In order to evade the *binding criterion* of the bit-commitment protocol (the requirement that after making the commitment to a bit value, it is not possible to change it during the opening phase of the protocol), Alice establishes perfect correlations between her and Bob's results for *one of the two different choices* of local measurement bases (computational for committing to 0, and diagonal for committing to 1), i.e. during the opening phase she informs Bob which measurement to perform. In our protocol, if using entanglement to cheat, the choice of pairs of local observables, for which perfect (or at least, better than random) correlations between Alice's and Bob's measurement results is achieved, is irrelevant. Thus, Alice performing local measurements on entangled states is equivalent to locally preparing at her site a simple statistical mixture of the corresponding collapsed Bob's states, and then sending them to him. Thus, the entanglement strategy is indeed equivalent to sending simple statistical mixture of non-entangled states. Below, we show that for every quantum state $|\psi_{\rm ch}\rangle$ of Bob's particle, the probability to obtain the intended message (and thus the probability of Alice's inference that Bob obtained the message) is bounded from above by a function that is exponentially close, with respect to the message length k, to 1/2.

Let l be the number of s_i's for which $\varphi_i = s_i\theta_n/2 \in [\pi/8; 3\pi/8] \cup [5\pi/8; 7\pi/8]$. For such cases we can consider the rearranged secret key $\mathbf{s} = s_1 \ldots s_l$ and the corresponding message $\mathbf{m} = m_1 \ldots m_l$. Depending on his choice of rotation direction a' Bob will measure one of the two observables $C_\pm(\mathbf{s}) = \sum_{\mathbf{m}=0}^{2^l-1} \mathbf{m} \times P_\pm(\mathbf{m};\mathbf{s})$, where one-dimensional projectors are given by $P_\pm(\mathbf{m};\mathbf{s}) = \bigotimes_{i=0}^{l} P_\pm(m_i;s_i) = \bigotimes_{i=0}^{l} |m_i(s_i)\rangle_\pm\langle m_i(s_i)|$ and messages (bit-strings) $\mathbf{m}$ are seen as binary numbers.

For given $\mathbf{m}$ and $\mathbf{s}$ Alice wants to maximize the probability $\Pr_{\rm ch}$ of Bob obtaining $\mathbf{m}$ measuring $C_\pm(\mathbf{s})$ on $|\psi_{\rm ch}\rangle$ (and thus her probability of knowing if he got the message or not), which is given by

$$\Pr_{\rm ch} = \frac{1}{2}\left(\|P_+(\mathbf{s})|\psi_{\rm ch}\rangle\|^2 + \|P_-(\mathbf{s})|\psi_{\rm ch}\rangle\|^2\right). \tag{20}$$

The state $|\psi_{\rm ch}\rangle$ that maximizes the above expression is the equal superposition of $|+\rangle = \prod_{i=1}^{l} |m_i(s_i)\rangle_+$ and $|-\rangle = \prod_{i=1}^{l} |m_i(s_i)\rangle_-$. Thus, we have:

$$\Pr_{\rm ch} \leqslant \frac{1}{2}\left(1 + |\langle +|-\rangle|\right) \leqslant \frac{1}{2}\left(1 + \cos^l(\pi/8)\right). \tag{21}$$

If the values s_i were produced uniformly at random, then the probability that $\varphi_i = s_i\theta_n/2 \in [\pi/8; 3\pi/8] \cup [5\pi/8; 7\pi/8]$ is 1/2. As a consequence, the random variable that

counts the number l of such φ_i's follow the binomial distribution $\mathcal{B}(k, 1/2)$, with k being the number of trials (the total number of rotation angles φ_i, equal to the length of the message $\mathbf{m}$) and 1/2 being the success probability of each trial (where by 'success' we mean that the rotation angle falls within the above mentioned intervals). For sufficiently large k, the binomial distribution can be approximated by the normal distribution $\mathcal{N}(\mu, \sigma^2)$ with the mean $\mu = k/2$ and the variance $\sigma^2 = k/4$. This allows Bob to set the degree of confidence of Alice's obliviousness. For example, choosing the 3σ criterion, if $(k - 3\sqrt{k})/2 \leqslant l \leqslant (k + 3\sqrt{k})/2$ Alice's probability to learn if Bob got the message or not will be $\mathrm{Pr}_{\mathrm{ch}} = 1/2 + \varepsilon(k)$, where $\varepsilon(k)$ is negligible (which happens in 99.8% of the cases if s_i were chosen uniformly at random).

Regarding the quantitative evaluation of the security levels against the cheating agents in terms of θ_n and k, we note that one can perform the estimation analogous to the one performed for the soundness criterion, presented above in section 3.1. Indeed, the security levels, given through the cheating probabilities, are for the probabilistic transfer and the (single-qubit measurements) obliviousness criteria given by (17) and (20), respectively, which have the similar dependence on the values $\cos(s_i\theta_n)$, for a given $\mathbf{s}$, to that of $\varepsilon(k)$ from (15). Regarding the concealingness criterion presented in section 3.2, while estimating its numerical security in terms of θ_n and k is, for general coherent multi-qubit measurements, a non-trivial task, one can straightforwardly see that the numerical expressions for the security levels for the case of single- and few-qubit measurements would again involve expressions similar the the ones considered above, given in terms of $\cos(s_i\theta_n)$.

Finally, we note that, according to the security criterion adopted in this paper, our OT protocol is secure against violating *only one* out of those three criteria (concealingness, probabilistic transfer and obliviousness), while *keeping the other two satisfied*. In case a cheating Alice decides not to send a message at all (by sending a wrong secret key s), she knows with certainty that Bob will not receive it (the protocol is not oblivious), at the same time violating the probabilistic requirement. This compromises the mentioned Crépeau's reduction [6] to a 1-out-of-2 OT, as the cheating Alice would know with probability 1/2 which of the two messages Bob chose to receive [55]. Nevertheless, using the reduction presented in [56] and a bit-commitment protocol (such as those presented in [27, 28, 34, 35]), one can achieve a 1-out-of-2 OT through the Crépeau's reduction based on a protocol such as ours (see [57] for a detailed discussion on the example of the computationally secure OT presented in [38]). Moreover, one can use our protocol to achieve two-party computation even without the help of a bit commitment. The semi-honest 1-out-of-2 OT based on our OT protocol is insecure against a cheating Alice. Due to the symmetry of the OT protocol [58], it can be further transformed into a semi-honest 1-out-of-2 OT insecure against a cheating Bob, who would with probability 1/2 obtain both messages sent by Alice (and only one in the remaining half of the cases). This is nothing but the 'cut-and-choose OT primitive', introduced in [59] and further developed in [60], which allows for two-party computation (for cut-and-choose protocols on garbled circuits [12], see also [61–63]).

Note that the security of the complex multi-party protocols is at most as high as the security of their building blocks: for a complex scheme to have certain level of security, its building blocks have to maintain the same security level, but this is not enough—a complex scheme could be insecure against, say, coherent measurements performed over all of its building blocks, etc. Our protocol is proven secure only against 'few-qubit measurements', i.e. it is only practically secure, its security being guaranteed by the technological limitations. Thus, our reduction to a single-bit OT, as well as the above mentioned application [59] to achieve multi-party computation without the use of BC, are proven to be only practically, and *not unconditionally*, secure. Note that the latter is a classical reduction, thus further

compromising its security which, according to Lo–Chau–Mayers no-go results, cannot be unconditional. As a consequence, no matter how high the security level our bit-string OT has, this fact *alone* is not in contradiction with Lo–Chao–Mayers no-go theorems.

In case it turns out that our protocol is indeed insecure against general coherent attacks, such result would support the impossibility of unconditionally secure single-bit all-or-nothing OT (anticipated by the significant portion of the community), and be in tune with the proven no-go theorems, in connection to the above-mentioned multi-party computation scheme [59]. Nevertheless, from the technical point of view, it would be interesting to see the particular coherent attack.

In case it turns out that our protocol is secure against general coherent attacks, it would not directly contradict no-go theorems, as our proposal is a bit-string all-or-nothing OT which, similarly to a bit-string commitment, in quantum domain does not necessarily have to be equivalent to a single-bit counterpart. Nevertheless, the question of the security level of our reduction from a bit-string to a single-bit all-or-nothing OT becomes an interesting one. In this case, two possible outcomes are: (i) similarly to the case of the application to multi-party computation [59], our reduction is insecure against coherent attacks (for example, proving the security against learning the string of bits does not necessarily mean that a cheating Bob cannot learn its parity instead, thus breaking the reduction security), or (ii) our single-bit all-or-nothing OT is secure against coherent attacks. Note that, while the predominant opinion within the community is that such security level is impossible, there exists an opposite controversial result [40], accompanied by the concrete construction of allegedly secure all-or-nothing OT [39]. Despite the recent proposal of the reduction from all-or-nothing to 1-out-of-2 OT [43], there exists no broadly accepted clearcut proof of the equivalence of the two flavours of the single-bit OT.

To summarise, our result is not in contradiction to any no-go theorem, and while our bit-string OT protocol is probably insecure against coherent attacks, even the opposite is allowed by the known impossibility proofs.

4. Conclusions

In this paper we proposed a novel scheme for obliviously transferring a bit-string message from Alice to Bob. The scheme presented does not violate the Lo's no-go theorem [30] and its security is based on the laws of quantum physics.

We proved that the protocol is secure against *any* cheating strategy of Bob before the opening phase (the protocol is concealing) as well as against a cheating Alice (it is oblivious, providing Alice maintains the probabilistic transfer). Furthermore, we proved that it satisfies probabilistic transfer, provided Bob performs only 'few-qubit' measurements, bounded by a certain upper bound ℓ (the protocol is practically probabilistic). Although intuitively our protocol should, at least for sufficiently large n, be secure against multi-qubit measurements, a detailed analysis of its security against Bob's coherent attacks remains to be done (similarly as for the case of recently proposed and performed quantum signature protocols [64, 65]).

Our protocol does not use entanglement and its optical implementation could be performed using today's technology. Indeed, as already mentioned, the current technology can perform single-photon polarisation rotations for the values on $n = 10$ (and possibly higher), which, using the standard photon emission and detection techniques (used, for example, in quantum key distribution) can achieve OT protocols with short time differences between the transferring and the opening phases. To achieve longer time differences, one would need to use stable long-term quantum memories, which are beyond the current technology. Nevertheless, even

the OT protocols with short times between the two phases have a number of useful examples of multi-party computation applications, such as private data mining, zero-knowledge proofs, etc. With the future advances of the technology, one can expect significant improvements in producing stable quantum memories which could, potentially, allow for our protocol to be used in applications that require for longer time-differences, such as quantum e-voting.

Finally, we discuss the need for the use of hash functions. Recall that at the end of the protocol Bob must be sure if he got the intended message or not. This property is guaranteed by comparing the computed hash value of the received message $\mathbf{m}$ with the presumed hash value sent by Alice together with $\mathbf{m}$. Such acknowledgment of the validity of the message decoded by Bob could be done differently. Suppose that out of all possible messages (PM), Alice is constrained to send $\mathbf{m}$ from a smaller set of messages (VM), such that verifying that $\mathbf{m}$ is in VM can be easily done, but only Alice knows the elements of VM. Note that in order to keep the probability of receiving a message from Alice to 1/2, up to a negligible term, the size of VM must be exponentially smaller than the size of PM. For example, VM could be the set of solutions to a hard mathematical problem, say 3-SAT problem. Alternatively, the message sent might be written in an existing human language, say English, making it easily recognisable by any English-language speaker.

Apart from (dis)proving the concealingness security criterion against the general coherent attacks (performing a measurement on all qubits at once), future lines of research include formulating other quantum security protocols that use single-qubit rotations to encode bit values into quantum states taken from a number of different bases. One such immediate application is in designing a quantum bit-string commitment protocol and compare it with the existing proposals. Furthermore, similarly when generating (randomised) secret keys, single-qubit rotations could be used in creating undeniable signatures.

In the proof of obliviousness of the protocol, section 3.4, we considered as relevant the encoded basis for which $\phi_i = s_i\theta_n/2 \in [\pi/8, 3\pi/8] \cup [5\pi/8, 7\pi/8]$. This might suggest that one can modify the protocol by using only those angles as possible encryption bases. Note though that the concealingness of the protocol is based in the security of the public-key cryptosystem of [44], which requires the whole range of equidistant angles. We leave this as an interesting question for future research. Another interesting possibility would be to further restrict the angles used, to those corresponding only to the computational and the diagonal bases, thus sending only BB84 states. Note though that such protocol would differ from our original idea in one important issue, that might affect its security. Namely, for the choice of the states from the computational bases, the cases of $a = 0$ and $a = 1$ would be indistinguishable (indeed was one of the reasons to exclude the case $\phi \approx 0$ from consideration in the proof of obliviousness of our protocol). We leave this interesting possibility as a topic for a future research.

An important future line of research is to go beyond the 'proof of principle' presented in this paper, and analyse quantitative effects of imperfect sources, noise and measurement errors in the protocol's optical realisations. In performing such study of the protocol's quantitative security levels in realistic applications, one can straightforwardly generalise the techniques used in the study of noise and measurement errors of a practical two- and four-state bit-commitment protocols, presented in [35, 66].

Acknowledgments

The authors acknowledge the support of SQIG—Security and Quantum Information Group, the Instituto de Telecomunicações (IT) Research Unit, ref. UID/EEA/50008/2013, the IT project QbigD funded by Fundação para a Ciência e Tecnologia (FCT) PEst-OE/EEI/LA0008/2013,

and the FCT project Confident PTDC/EEI-CTP/4503/2014. A.S. also acknowledges the FCT Post-doc scholarship SFRH/BPD/76231/2011 during which the major part of the work was done and LaSIGE Research Unit, ref. UID/CEC/00408/2013.

Appendix A. Authentication in multiparty computation

Oblivious transfer is a building block used to construct secure multi-party computation. Loosely speaking, the secure multi-party computation of a function $g(i_1, \ldots, i_k) = (o_1, \ldots, o_k)$ is a protocol among k agents, which do not trust each other. The aim of the protocol is to jointly compute function g such that, each agent A_ℓ, with $\ell \in \{1 \ldots k\}$:

- inputs its private message i_ℓ, which is kept secret to the other agents;
- receives the private output o_ℓ, known only by A_ℓ.

An example of a secure multi-party computation is private data mining, where there are two agents ($k = 2$), say Alice and Bob, and Alice wishes to perform some statistics s over a private database d of Bob. Thus, the function to compute is of the form $g(\perp, d) = (s(d), \perp)$, where $\perp$ means the empty string.

If Bob requires Alice to authenticate herself, in order to perform the statistics, a simple way to enforce the authentication is for Alice to input a password w, and the statistics will only be given if the password is correct, that is, if the hash of the password coincides with the stored password hash $h_A = h(w)$ of Alice in Bob's server (similar to computer login),

$$g(w, (d, h_A)) = \begin{cases} (s(d), \perp) & \text{if } h(w) = h_A \\ (\perp, \perp) & \text{otherwise.} \end{cases} \tag{A.1}$$

This idea can be generalised to any secure multi-party computation, assuming that each party shares the hashes of the passwords of each all other agents. And so, instead of computing the function $g(i_1, \ldots, i_k) = (o_1, \ldots, o_k)$, one computes the function

$$g((i_1, w_1, \mathbf{h}_1), \ldots, (i_k, w_k, \mathbf{h}_k)) = \begin{cases} (o_1, \ldots, o_k) & \text{if } h(w_1) = h_{A_1} \wedge \cdots \wedge h(w_k) = h_{A_k} \\ (\perp, \ldots, \perp) & \text{otherwise,} \end{cases} \tag{A.2}$$

where $\mathbf{h}_\ell = (h_{A_1}, \ldots, h_{A_{\ell-1}}, h_{A_{\ell+1}}, \ldots, h_{A_k})$ is the string of hashes of the passwords of all other agents but A_ℓ. More details can be found in [45].

Appendix B. Notation, definitions and results

In this appendix, we provide notation, necessary definitions and results for stating and proving the security of our proposal. First, we give the definition of *quantum one-way functions*, based on [44, 67].

A *quantum one-way function* is a map $f : N \to \mathcal{H}$, where $N \subset \mathbb{Z}$ and $\mathcal{H}$ is a Hilbert space, such that:

(i) f is easy to compute: there is a polynomial-time (in the number of bits of the input s) quantum algorithm that computes $f(s) \in \mathcal{H}$, with $s \in N$;
(ii) Hard to invert: without any additional information, inverting $f(s)$ non-negligibly on the input size is impossible by fundamental physical laws of quantum (information) theory.

If inversion is computed in polynomial-time using some additional information, called the trapdoor, then f is called a *trapdoor quantum one-way function*. In a sense, the trapdoor is a 'key' to unlock the input s.

There are several candidates for quantum one-way functions studied in [68, 69] (a slightly different variation of quantum one-way function, with input being quantum as well, was considered in [67]). Recently, another candidate for a quantum one-way function was proposed in [44]. This function considers qubit rotations R and is given by

$$f(s) = R(s\theta_n)|0\rangle = \cos(s\theta_n/2)\,|0\rangle + \sin(s\theta_n/2)\,|1\rangle \tag{B.1}$$

where $s \in \{0, \ldots, 2^n - 1\}$ and $\theta_n = \pi/2^{n-1}$, for some fixed n, and $\{|0\rangle, |1\rangle\}$ is a fixed computational basis (i.e. f is not a function of a quantum state). Notice that this is a quantum one-way function because:

- Qubit rotations $R(s\theta_n)$ can be easily implemented up to an arbitrary accuracy by a quantum algorithm involving an universal set of gates ([70, 44]).
- Due to Holevo bound, the maximal amount of information that can be extracted by means of a POVM on a single qubit is 1 bit. Since s has n bits, it is impossible to recover s from a single qubit in the state $R(s\theta_n)|0\rangle$.

Moreover, f can be used to construct a quantum trapdoor one-way function $F(s, b)$, where s is the trapdoor information for learning an unknown bit b [44]:

$$F(s, b) = R(b\pi)f(s) = R(b\pi)R(s\theta_n)|0\rangle = R(s\theta_n + b\pi)|0\rangle. \tag{B.2}$$

Note that inverting F (learning both s and b) is at least as hard as inverting f. Also, the ensemble of qubits, each in a state $F(s_i, b_i)$, where s_i and b_i are random, is described by a complete mixture $\rho = \mathbb{1}/2$, if s_i and b_i are unknown [44]. Therefore, every binary measurement that could be used to infer unknown bit b would give completely random value. Nevertheless, if s is known, by applying the rotation $R(-s\theta_n)$ to $F(s, b)$ and measuring the result in the computational basis, one obtains b with certainty. Therefore, $F(s, b)$ is a polynomial quantum trapdoor one-way function.

Based on the above discussion, we present the secure public-key cryptosystem proposed in [44]:

Protocol 3 (Public-key encryption scheme).

Message to transfer: $\mathbf{m} = m_1 \ldots m_l$ *with* $l \leqslant k$;
Security parameter: n;
Secret key: $\mathbf{s} = (s_1, \ldots, s_k)$, *where each* $s_i \in \{0, \ldots, 2^n - 1\}$;

Public key generation:

(i) *For all* $1 \leqslant i \leqslant k$, *Alice chooses uniformly at random* $s_i \in \{0, \ldots, 2^n - 1\}$, *and* $\mathbf{s} = (s_1, \ldots, s_k)$ *will be her private key.*
(ii) *Alice generates the corresponding public key:*

$$|\psi\rangle = \bigotimes_{i=1}^{k} R(s_i\theta_n)|0\rangle \tag{B.3}$$

$$= \bigotimes_{i=1}^{k} \left(\cos\left(\frac{s_i\theta_n}{2}\right)|0\rangle + \sin\left(\frac{s_i\theta_n}{2}\right)|1\rangle\right). \tag{B.4}$$

Encryption:

(iii) *Bob wishes to send message* $\mathbf{m} = m_1, \ldots m_l$ *where* $l \leqslant k$.
(iv) *Bob obtains Alice's public key,* $|\psi\rangle$.

(v) *Bob encrypts his message* $\mathbf{m}$ *(padded with 0 if necessary) as follows*

$$|\psi(\mathbf{m})\rangle = \bigotimes_{i=1}^{k} R(m_i\pi)|\psi\rangle \tag{B.5}$$

$$= \bigotimes_{i=1}^{k} \left(\cos\left(\frac{s_i\theta_n}{2} + \frac{m_i\pi}{2}\right)|0\rangle + \sin\left(\frac{s_i\theta_n}{2} + \frac{m_i\pi}{2}\right)|1\rangle\right). \tag{B.6}$$

(vi) *Bob sends* $|\psi(\mathbf{m})\rangle$ *to Alice.*

Decryption:

(vii) *Alice uses private key as follows*

$$|\psi'(\mathbf{m})\rangle = \bigotimes_{i=1}^{k} R(-s_i\theta_n)|\psi(\mathbf{m})\rangle \tag{B.7}$$

$$= \bigotimes_{i=1}^{k} \left(\cos\left(\frac{m_i\pi}{2}\right)|0\rangle + \sin\left(\frac{m_i\pi}{2}\right)|1\rangle\right) \tag{B.8}$$

$$= \bigotimes_{i=1}^{k} |m_i\rangle. \tag{B.9}$$

(viii) Alice performs measurements on each $|m_i\rangle$ *in the computational basis.*

Obviously, the public-key generation corresponds to the computation of $f(s)$, the encryption phase computes $F(s, b)$ and the decryption phase corresponds to the inversion of $F(s, b)$ with the trapdoor information s, which allows to learn message $\mathbf{m}$. As discussed in sections IV.B and IV.C of [44] and also in [51], even if Alice publicly announces n the system is still secure. In fact, using the Holevo bound, the cryptographic scheme is information secure if k, the length of the message is of same order than n. In fact, the maximum information that Bob can obtain from the public key regarding the secret key is k bits, but its uncertainty is nk, which in case of $n = k$ implies a negligible advantage for an adversary to recover the message. Furthermore, since any rotation leaves a complete mixed state invariant, all possible messages yield the same cipher state, which is equal to a public-key state, and therefore adversaries cannot distinguish distinct messages, unless many copies of the encrypted state and the respective public key are provided, making the scheme provably secure [71, 72].

In figure B1 we present a schematic description of the public-key cryptosystem.

At the end of the Oblivious transfer protocol, Bob has to be assured if he received the message or not. There are different ways to guarantee this feature. The solution adopted in this paper is to use a *hash function*. A hash function maps strings to other strings of smaller size. Therefore, different strings are mapped to the same hash value. Hash functions have to satisfy the following two constrains,

- their value for each input can be computed in polynomial time on the length of the input string;
- the hash values of a randomly chosen string are uniformly distributed.

Consider two sets A and B of size a and b, respectively, such that $a > b$, and consider a hash function $h : A \to B$. It is easy to derive that the size of a set $A_{\mathbf{x}}$ of strings $\mathbf{x} \in A$ mapped to the same hash value $h(\mathbf{x})$ is at most a/b.

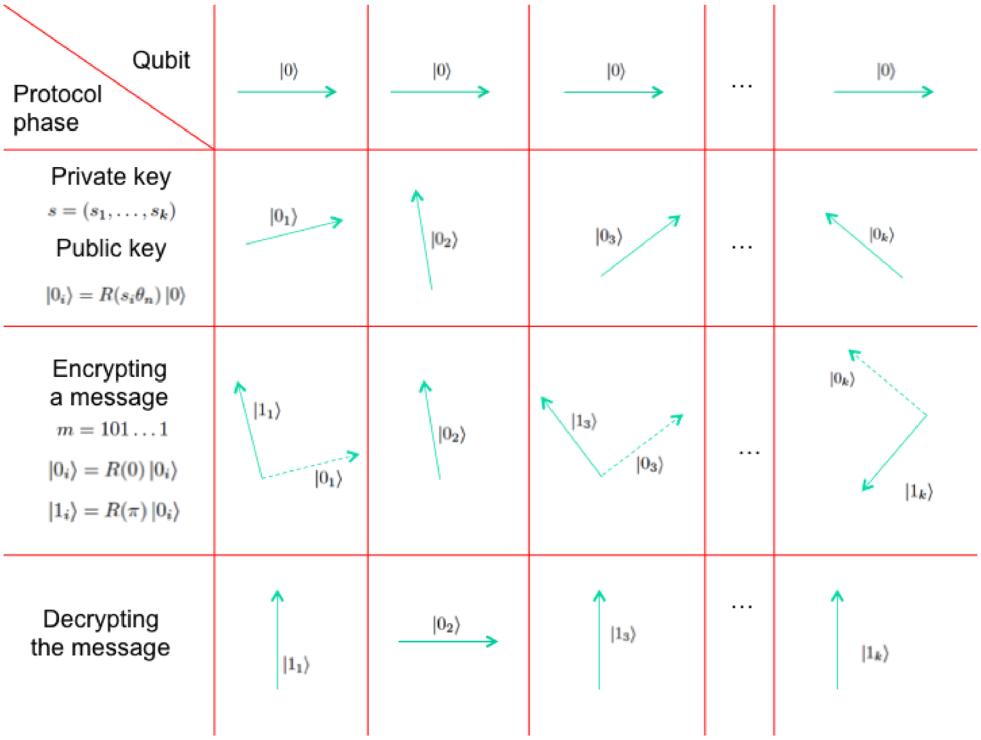


Figure B1. Schematic description of the public key cryptosystem for messages of length k. The full arrows represent the actual states of qubits, while the dashed arrows in the third line (encryption of a message) represent $|0_i\rangle$ states.

In particular, requiring that A contains all strings of length k and B to be a set of strings of length $\omega = \lfloor\sqrt{k}\rfloor$, the number of strings with the same hash value is 2^{ω}, hence the probability of finding such a string is negligible in k. For more details on constructing hash functions, see for example [73].

References

[1] Bennett C and Brassard G 1984 Quantum cryptography: public key distribution and coin tossing *Proc. of the IEEE Int. Conf. on Computers, Systems and Signal Processing* (New York: IEEE) pp 175–9
[2] Ekert A 1991 Quantum cryptography based on bell's theorem *Phys. Rev. Lett.* **67** 661–3
[3] Bennett C 1992 Quantum cryptography using any two nonorthogonal states *Phys. Rev. Lett.* **68** 3121–4
[4] Rabin M 1981 How to exchange secrets by oblivious transfer *Technical Report* TR-81 (Harvard Aiken Computation Laboratory)
[5] Even S, Goldreich O and Lempel A 1985 A randomized protocol for signing contracts *Commun. ACM* **28** 637–47
[6] Crépeau C 1988 Equivalence between two flavours of oblivious transfers *A Conf. on the Theory and Applications of Cryptographic Techniques on Advances in Cryptology* (London: Springer) pp 350–4
[7] Brassard G, Crépeau C and Robert J 1986 Information theoretic reductions among disclosure problems *27th Annual Symp. on Foundations of Computer Science* pp 168–73

[8] Crépeau C and Santha M 1993 Efficient reduction among oblivious transfer protocols based on new self-intersecting codes *Sequences II* ed R Capocelli *et al* (New York: Springer) pp 360–8
[9] Brassard G, Crépeau C and Santha M 1996 Oblivious transfers and intersecting codes *IEEE Trans. Inf. Theory* **42** 1769–80
[10] Kilian J 1988 Founding crytpography on oblivious transfer *Proc. of the 20th Annual ACM Symp. on Theory of Computing* (New York: ACM) pp 20–31
[11] Harn L and Lin H 1993 An oblivious transfer protocol and its application for the exchange of secrets *Advances in Cryptology* (*Lecture Notes in Computer Science* vol 739) ed H Imai *et al* (Berlin: Springer) pp 312–20
[12] Yao A 1986 How to generate and exchange secrets *Proc. of the 27th Annual Symp. on Foundations of Computer Science* (Washington, DC: IEEE Computer Society) pp 162–7
[13] Cramer R, Damgård I and Maurer U 2000 General secure multi-party computation from any linear secret-sharing scheme *Proc. of the 19th Int. Conf. on Theory and Application of Cryptographic Techniques* (Berlin: Springer) pp 316–34
[14] Lindell Y and Pinkas B 2012 Secure two-party computation via cut-and-choose oblivious transfer *J. Cryptol.* **25** 680–722
[15] Lindell Y and Zarosim H 2013 On the feasibility of extending oblivious transfer *Theory of Cryptography Conf.* (Berlin: Springer) pp 519–38
[16] Brassard G, Chaum D and Crépeau C 1988 Minimum disclosure proofs of knowledge *J. Comput. Syst. Sci.* **37** 156–89
[17] Salvail L 1999 *The Search for the Holy Grail in Quantum Cryptography* (Berlin: Springer) pp 183–216
[18] Bennett C, Brassard G, Crépeau C and Skubiszewska M 1992 Practical quantum oblivious transfer *Advances in Cryptology* (*Lecture Notes in Computer Science* vol 576) ed J Feigenbaum (Berlin: Springer) pp 351–66
[19] Wiesner S 1983 Conjugate coding *SIGACT News* **15** 78–88
[20] Lo H K and Chau H 1999 Unconditional security of quantum key distribution over arbitrarily long distances *Science* **283** 2050–6
[21] Shor P W and Preskill J 2000 Simple proof of security of the bb84 quantum key distribution protocol *Phys. Rev. Lett.* **85** 441–4
[22] Mayers D 2001 Unconditional security in quantum cryptography *J. ACM* **48** 351–406
[23] Scarani V, Bechmann-Pasquinucci H, Cerf N, Dušek M, Lütkenhaus N and Peev M 2009 The security of practical quantum key distribution *Rev. Mod. Phys.* **81** 1301–50
[24] Paunković N, Bouda J and Mateus P 2011 Fair and optimistic quantum contract signing *Phys. Rev.* A **84** 062331
[25] Lo H K and Chau H F 1997 Is quantum bit commitment really possible? *Phys. Rev. Lett.* **78** 3410–3
[26] Mayers D 1997 Unconditionally secure quantum bit commitment is impossible *Phys. Rev. Lett.* page 3414
[27] Kent A 1999 Unconditionally secure bit commitment *Phys. Rev. Lett.* **83** 1447–50
[28] Kent A 2005 Secure classical bit commitment using fixed capacity communication channels *J. Cryptol.* **18** 313–35
[29] Bouman N, Fehr S, Gonzalez-Guillen C and Schaffner C 2013 An all-but-one entropic uncertainty relation and application to password-based identification *Theory of Quantum Computation, Communication and Cryptography* (*Lecture Notes in Computer Science* vol 7582) ed K Iwama *et al* (Berlin: Springer) pp 29–44
[30] Lo H K 1997 Insecurity of quantum secure computations *Phys. Rev.* A **56** 1154–62
[31] Wehner S, Schaffner C and Terhal B M 2008 Cryptography from noisy storage *Phys. Rev. Lett.* **100** 220502
[32] Schaffner C, Terhal B M and Wehner S 2011 Robust cryptography in the noisy-quantum-storage model *Quantum Inf. Comput.* **9** 963–96
[33] Koenig R, Wehner S and Wullschleger J 2012 Unconditional security from noisy quantum storage *IEEE Trans. Inf. Theory* **58** 1962–84
[34] Ng N, Joshi S, Ming C, Kurtsiefer C and Wehner S 2012 Experimental implementation of bit commitment in the noisy-storage model *Nat. Commun.* **3** 1326
[35] Loura R, Almeida Á, André P, Pinto A, Mateus P and Paunković N 2014 Noise and measurement errors in a practical two-state quantum bit commitment protocol *Phys. Rev.* A **89** 052336
[36] Almeida Á, Loura R, Paunković N, Silva N, Muga N, Mateus P, André P and Pinto A 2014 A brief review on quantum bit commitment *Proc. SPIE* **9286** 92861C

[37] Almeida Á, Stojanovic A, Paunković N, Loura R, Muga N J, Silva N, Mateus P, André P and Pinto A 2016 Implementation of a two-state quantum bit commitment protocol in optical fibers *J. Opt.* **18** 015202
[38] Souto A, Mateus P, Adão P and Paunković N 2015 Bit-string oblivious transfer based on quantum state computational distinguishability *Phys. Rev.* A **91** 042306
[39] He G P and Wang Z D 2006 Oblivious transfer using quantum entanglement *Phys. Rev.* A **73** 012331
[40] He G P and Wang Z D 2006 Nonequivalence of two flavors of oblivious transfer at the quantum level *Phys. Rev.* A **73** 044304
[41] Kent A 2003 Quantum bit string commitment *Phys. Rev. Lett.* **90** 237901
[42] Barrett J and Massar S 2004 Security of quantum bit-string generation *Phys. Rev.* A **70** 052310
[43] Pivoluska M, Plesch M and Pawlowski M 2016 1-out-of-2 oblivious transfer using flawed bitstring quantum protocol (arXiv:1611.10087)
[44] Nikolopoulos G 2008 Applications of single-qubit rotations in quantum public-key cryptography *Phys. Rev.* A **77** 032348
[45] Goldreich O 2002 Secure Multiparty Computation (v1.4 Final Draft) Weizmann Institute of Science (http://www.wisdom.weizmann.ac.il/~oded/PS/prot.ps)
[46] Canetti R, Dachman-Soled D, Vaikuntanathan V and Wee H 2012 Efficient password authenticated key exchange via oblivious transfer *Proc. of PKC 2012—15th Int. Conf. on Practice and Theory in Public Key Cryptography (Darmstadt, Germany, 21–23 May 2012)* (*Lecture Notes in Computer Science* vol 7293) (Berlin: Springer) pp 449–66
[47] Jain R 1991 *The Art of Computer Systems Performance Analysis—Techniques for Experimental Design, Measurement, Simulation and Modeling* (New York: Wiley)
[48] Goldreich O 2004 *Foundations of Cryptography: Volume 2, Basic Applications* (New York: Cambridge University Press)
[49] Black P 2016 'Polylogarithmic' *Dictionary of Algorithms and Data Structures* ed V Pieterse and P E Black (Washington, DC: NIST) www.nist.gov/dads/HTML/rootedtree.html (accessed 1 February 2016)
[50] Vlachou C, Rodrigues J, Mateus P, Paunković N and Souto A 2015 Quantum walk public-key cryptographic system *Int. J. Quantum Inf.* **13** 1550050
[51] Seyfarth U, Nikolopoulos G and Alber G 2012 Symmetries and security of a quantum-public-key encryption based on single-qubit rotations *Phys. Rev.* A **85** 022342
[52] 2017 Thor Labs New Jersey, USA *Polarisation Instrumentation* (http://www.wisdom.weizmann.ac.il/~oded/PS/prot.ps)
[53] Helstrom C 1969 Quantum detection and estimation theory *J. Stat. Phys.* **1** 231–52
[54] Mayers D 1995 On the security of the quantum oblivious transfer and key distribution protocols *CRYPTO* (*Lecture Notes in Computer Science* vol 963) ed D Coppersmith (Berlin: Springer) pp 124–35
[55] He G P 2015 Comment on 'Bit-string oblivious transfer based on quantum state computational distinguishability' *Phys. Rev.* A **92** 046301
[56] Choi S, Dachman-Soled D, Malkin T and Wee H 2009 *Simple, Black-Box Constructions of Adaptively Secure Protocols* (Berlin: Springer) pp 387–402
[57] Souto A, Mateus P, Adão P and Paunković N 2015 Reply to 'Comment on "Bit-string oblivious transfer based on quantum state computational distinguishability" ' *Phys. Rev.* A **92** 046302
[58] Wolf S and Wullschleger J 2006 Oblivious transfer is symmetric *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (Berlin: Springer) pp 222–32
[59] Lindell Y and Pinkas B 2011 *Secure Two-Party Computation via Cut-and-Choose Oblivious Transfer* (Berlin: Springer) pp 329–46
[60] Lindell Y 2016 Fast cut-and-choose-based protocols for malicious and covert adversaries *J. Cryptol.* **29** 456–90
[61] Lindell Y and Pinkas B 2007 An efficient protocol for secure two-party computation in the presence of malicious adversaries *Annual Int. Conf. on the Theory and Applications of Cryptographic Techniques* (Berlin: Springer) pp 52–78
[62] Mohassel P and Riva B 2013 Garbled circuits checking garbled circuits: More efficient and secure two-party computation *Annual Cryptology Conf.* (Berlin: Springer) pp 36–53
[63] Shelat A and Shen C H 2011 Two-output secure computation with malicious adversaries *Annual Int. Conf. on the Theory and Applications of Cryptographic Techniques* (Berlin: Springer) pp 386–405

[64] Dunjko V, Wallden P and Andersson E 2014 Quantum digital signatures without quantum memory *Phys. Rev. Lett.* **112** 040502
[65] Collins R, Donaldson R, Dunjko V, Wallden P, Clarke P, Andersson E, Jeffers J and Buller G 2014 Realization of quantum digital signatures without the requirement of quantum memory *Phys. Rev. Lett.* **113** 040502
[66] Loura R, Arsenović D, Paunković N, Popović D B and Prvanović S 2016 Security of two-state and four-state practical quantum bit-commitment protocols *Phys. Rev.* A **94** 062335
[67] Lu X and Feng D 2005 Quantum digital signature based on quantum one-way functions *The 7th Int. Conf. on Advanced Communication Technology* vol 1 pp 514–7
[68] Gottesman D and Chuang I 2001 Quantum digital signatures *Technical Report* (arXiv:0105032)
[69] Buhrman H, Cleve R, Watrous J and De Wolf R 2001 Quantum fingerprinting *Phys. Rev. Lett.* **87** 2001
[70] Nielsen M and Chuang I 2004 *Quantum Computation and Quantum Information* (*Cambridge Series on Information and the Natural Sciences*) 1st edn (Cambridge: Cambridge University Press)
[71] Boykin P and Roychowdhury V 2003 Optimal encryption of quantum bits *Phys. Rev.* A **67** 042317
[72] Zou X and Qiu D 2013 Comment on 'Optimal encryption of quantum bits' *Phys. Rev.* A **88** 026302
[73] Carter J and Wegman M 1979 Universal classes of hash functions *J. Comput. Syst. Sci.* **18** 143–54

Security of two-state and four-state practical quantum bit-commitment protocols

Ricardo Loura,[1,2] Dušan Arsenović,[3] Nikola Paunković,[1,2,*] Duška B. Popović,[3] and Slobodan Prvanović[3]
[1]*Instituto de Telecomunicações, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal*
[2]*Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal*
[3]*Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia*
(Received 27 September 2016; published 27 December 2016)

We study cheating strategies against a practical four-state quantum bit-commitment protocol [A. Danan and L. Vaidman, Quant. Info. Proc. **11**, 769 (2012)] and its two-state variant [R. Loura *et al.*, Phys. Rev. A **89**, 052336 (2014)] when the underlying quantum channels are noisy and the cheating party is constrained to using single-qubit measurements only. We show that simply inferring the transmitted photons' states by using the Breidbart basis, optimal for ambiguous (minimum-error) state discrimination, does not directly produce an optimal cheating strategy for this bit-commitment protocol. We introduce a strategy, based on certain postmeasurement processes and show it to have better chances at cheating than the direct approach. We also study to what extent sending forged geographical coordinates helps a dishonest party in breaking the binding security requirement. Finally, we investigate the impact of imperfect single-photon sources in the protocols. Our study shows that, in terms of the resources used, the four-state protocol is advantageous over the two-state version. The analysis performed can be straightforwardly generalized to any finite-qubit measurement, with the same qualitative results.

DOI: 10.1103/PhysRevA.94.062335

I. INTRODUCTION

With the accelerating development of quantum information theory, new cryptographic protocols and primitives arise every day alongside new engineering techniques to turn them into reality. Naturally, due to the resemblance to classical cryptography and its implementations, new challenges and new difficulties appear once one moves from theory to practice. Quantifying the impact of these difficulties and studying their solutions is thus of utmost importance if quantum cryptography is to become a widespread reality. As a step towards this goal, we here study the feasibility of several different attacks on a practical two-state quantum bit-commitment (BC) protocol [1–3] and its original four-state variant [4] in a noisy environment. By practical, we mean that cheating is subjected to current technological constraints: the lack of long-term quantum memories and, in the case of the currently predominant optical realizations, the nonexistence of photon non-demolition measurements. Thus, an agent (in our case Alice) is forced to perform measurements as soon as he or she receives the qubits encoded in, say, photon polarization. Various practical cryptographic protocols based on bounded and/or noisy memories have been studied [5–8], in order to avoid the Lo-Chau [9] and Mayers [10] no-go theorem regarding nonrelativistic implementations of quantum bit-commitment schemes. (see also a recent proposal of a computationally secure oblivious transfer protocol [11]).

We start by reviewing the protocol's steps, including a careful definition of the verification procedure, in Sec. II. In Sec. III we discuss the so-called *cheating observable*, given by the Breidbart basis, and show that the naïve approach, studied in [3], can be perfected. In Section. IV, we analyze the so-called *faking-distance attack*, based on providing wrong geographical information to the verifier. Finally, Sec. V is dedicated to the so-called *beam-splitter attack*, which takes advantage of the fact that no perfect single-photon sources exist, much to the resemblance of BB84's photon-number splitting attack. We compare the enhanced attack with its counterpart in the four-state variant.

II. THE PROTOCOL

Commitment protocols, initially proposed in 1981 by Manuel Blum [12] and, simultaneously, by Shamir, Rivest and Adleman [13], are important current cryptographic primitives that serve as building blocks for several cryptographic and computational protocols, such as zero-knowledge proofs [14] and multiparty computation [15]. Simply put, a commitment scheme is a sequence of steps between a prover (Alice) and a verifier (Bob) that allows them to essentially roll a die in a fair way over a communication channel. In the first step (the *commitment phase*), Alice makes a choice from a set of possible choices $(1,\ldots,n$ in the case of an actual n-sided die). In the second step (the *revealing phase*), Bob asks Alice to reveal her choice and to prove that it was indeed her initial choice. In practice, these two steps are not done immediately after one another but are, instead, crucial parts of a larger protocol. Naturally, the scheme must guarantee that between the two phases Alice cannot change her mind (called the *binding* requirement) and that Bob cannot learn Alice's commitment choice before she reveals it (called the *concealing* requirement). For a more detailed description and examples of commitment schemes, see [3]. If Alice's choice is bound to be either 0 or 1, then we say that the commitment scheme is a bit-commitment scheme. Furthermore, if the protocol requires quantum communication of any kind, we say that it is a quantum commitment scheme.

In this work we are interested in a particular quantum bit-commitment scheme initially presented in [3], building on the earlier works of Danan and Vaidman in [4]. We begin with a brief review of the protocol's definition. A detailed version of what follows, along with a theoretical security analysis, can be found in [2,3]. Note, however, that in the present paper we

*npaunkov@math.tecnico.ulisboa.pt

 ©2016 American Physical Society

use a slightly different notation. In fact, the works [2,3] focus exclusively on the two-state protocol, whereas we here aim at studying both the two-state and the four-state protocol variants. Since the original notations of each of these are incompatible, we opt here for the standard quantum computation convention, recalled below.

The two-state protocol goes as follows (the analogous four-state protocol is presented in Sec. III B):

(1) Alice and Bob agree on a computational basis $\{|0\rangle,|1\rangle\}$ and, as usual, on the vectors $|+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$ and $|-\rangle = 1/\sqrt{2}(|0\rangle - |1\rangle)$.

(2) Bob randomly chooses one of the two states, $|0\rangle$ or $|+\rangle$, then prepares and sends a particle in that state to Alice. He repeats this step a number M of times of his choice.

(3) Alice either measures $\hat{C}_0$, for committing to 0, on all the particles she receives or measures $\hat{C}_1$, for committing to 1, on all the particles she receives, where $\hat{C}_0$ and $\hat{C}_1$ are the observables defined by

$$\begin{aligned}\hat{C}_0 &= 0\cdot|0\rangle\langle 0| + 1\cdot|1\rangle\langle 1|\\ \hat{C}_1 &= 0\cdot|+\rangle\langle +| + 1\cdot|-\rangle\langle -|.\end{aligned} \tag{1}$$

(4) Whenever Alice wants to reveal her commitment, she sends her commitment choice and the results of her measurements back to Bob.

(5) Bob either accepts or discards Alice's commitment, based on some statistical criterion of his choosing.

Since the states sent by Bob are nonorthogonal, the outcomes of Alice's measurements follow different probability distributions, according to whether $\hat{C}_0$ or $\hat{C}_1$ was the chosen observable. Indeed, if we denote by $p_i(j|k)$ the probability that Alice measures the value j with the observable $\hat{C}_i$ whenever Bob sends a particle in state $|k\rangle$, we have

$$\begin{aligned}p_0(0|0) &= 1,\\ p_0(1|0) &= 0,\\ p_0(0|+) &= \tfrac{1}{2},\\ p_0(1|+) &= \tfrac{1}{2}\end{aligned} \tag{2}$$

and, similarly,

$$\begin{aligned}p_1(0|0) &= \tfrac{1}{2},\\ p_1(1|0) &= \tfrac{1}{2},\\ p_1(0|+) &= 0,\\ p_1(1|+) &= 1.\end{aligned} \tag{3}$$

Choosing either observable $\hat{C}_0$ or observable $\hat{C}_1$ thus has a clear impact on Alice's set of measurement outcomes or, in other words, her sets of statistical data, her statistical signature. Consequently, once Alice reveals her results, Bob can know whether she used observable $\hat{C}_0$ or observable $\hat{C}_1$ (see below). Evidently, measuring any other observable, or different observables on each photon, will destroy the statistical signature. By exactly how much, and whether or not Alice can take any kind of advantage out of it, is the topic of the next section.

Unfortunately, things are not as simple in any real-world implementation of the protocol, where any emission, transmission, and measurement processes are inevitably subject to a certain amount of white noise. To account for such imperfections during the course of the protocol, we model the white noise as a depolarizing channel, or r-noise channel, $\mathcal{E}_d$, such that with probability r the state becomes totally mixed, while with probability $(1-r)$ it remains unharmed. It can thus be represented as a superoperator acting on a state ρ by

$$\mathcal{E}_d(\rho) = (1-r)\rho + r\frac{I}{2}. \tag{4}$$

As expected, such a channel affects the sets of probabilities (2) and (3), which now become functions of r. We abuse notation and refer to the new set of probabilities by the same symbols, as the former will not be of use henceforth. We have

$$\begin{aligned}p_0(0|0) &= 1-\frac{r}{2},\\ p_0(1|0) &= \frac{r}{2},\\ p_0(0|+) &= \frac{1}{2},\\ p_0(1|+) &= \frac{1}{2}\end{aligned} \tag{5}$$

and

$$\begin{aligned}p_1(0|0) &= \frac{1}{2},\\ p_1(1|0) &= \frac{1}{2},\\ p_1(0|+) &= \frac{r}{2},\\ p_1(1|+) &= 1-\frac{r}{2}.\end{aligned} \tag{6}$$

Finally, we need to formulate a systematic way for Bob to accept or discard Alice's commitment. Once Alice reveals her commitment and her measurement results, Bob's goal is to determine whether her statistics follow the probability distribution defined by (5), the one defined by (6), or neither of these two. This problem, of deciding whether or not a sample comes from a given distribution, is an extremely well-known problem in statistics, and can be solved through a variety of methods, jointly called goodness-of-fit tests, the most famous of which is likely to be the chi-square test. We here, however, take advantage of the simplicity of our particular case and instead use a far more straightforward test, called the binomial test.

To do so, let us assume that Bob sends N_0 particles in state $|0\rangle$. Without loss of generality, suppose Alice measures the $\hat{C}_0$ observable on all of them, i.e., she commits to the value 0. Then the probability $P(n_0\ 0\text{'s})$ that she obtains n_0 times the value 0 after measuring all the particles is given, by definition, by a binomial distribution:

$$P(n_0\ 0\text{'s}) = \binom{N_0}{n_0} p_0^{n_0}(0|0)\,p_0^{N_0-n_0}(1|0). \tag{7}$$

This binomial distribution has mean value μ_0 and variance σ_0^2 defined by

$$\begin{aligned}\mu_0 &= N_0 p_0(0|0),\\ \sigma_0^2 &= N_0 p_0(0|0)(1-p_0(0|0)).\end{aligned} \tag{8}$$

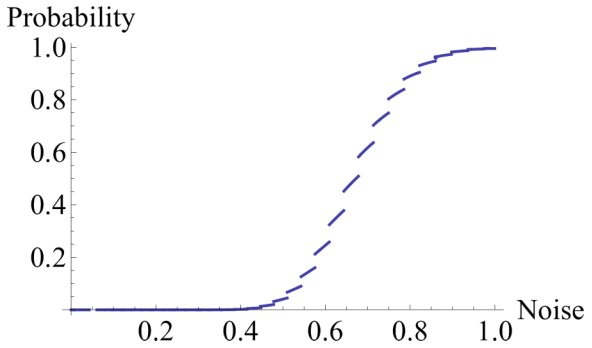


FIG. 1. Probability, as a function of white noise, of an honest Alice committing to 1, with $M = 100$ successful measurements, passing Bob's test of commitment to 0.

The binomial test then goes as follows: Bob proceeds if Alice's statistics satisfy

$$n_0 \in [\mu_0 - 3\sigma_0, \mu_0 + 3\sigma_0]. \tag{9}$$

Otherwise, he immediately aborts the protocol. The above condition ensures that if Alice is committing to 0, then she has ∼99.7% chance of having her first set of statistics accepted by Bob. This value is simply

$$0.997 \sim \sum_{k=\lceil \mu_0 - 3\sigma_0 \rceil}^{\lfloor \mu_0 + 3\sigma_0 \rfloor} \binom{N_0}{k} p_0^k(0|0) p_0^{N_0-k}(1|0). \tag{10}$$

Analogously, Bob tests Alice's statistics whenever state $|+\rangle$ is sent and proceeds if and only if

$$n_+ \in [\mu_+ - 3\sigma_+, \mu_+ + 3\sigma_+]. \tag{11}$$

Otherwise, he immediately aborts. The probability that Alice's commitment is accepted by Bob, i.e., the probability that Alice passes both tests, is then simply the product of the separate probabilities. Note, of course, that one could choose a standard deviation factor other than 3. The smaller it is, the more restrictive the criterion becomes. For simplicity, throughout the paper we set that Bob sends an equal number of particles in each state, i.e., $N_0 = N_+ = N$ (and $N_{0/1} = N_{+/-}$ in the four-state case), the assumption secured by the law of large numbers.

The test is, *mutatis mutandis*, the same if Alice were to commit to 1 instead (Fig. 1). However, there is still a loose end: there is a nonzero probability that an Alice trying to commit to 1 produces statistics that satisfy both (9) and (11), thus convincing Bob that she committed to 0 and making the test nonviable. This probability decreases exponentially with the number of measurements. Heuristically, a number of total measurements of $M = 100$ (about fifty 0's measured, and fifty 1's measured, $N = N_0 = N_+ = 50$) is enough to guarantee viability of the protocol for any reasonable amount of noise, as Fig. 1 shows.

III. OPTIMAL CHEATING OBSERVABLE

The main goal of this section is to answer a simple question: Can Alice cheat? In other words, can she devise a strategy so that she is able to convince Bob that she made a choice when in reality she did not? Naturally, Bob could also try to break the protocol. However, he has a much tougher job. For Bob to break the protocol, he would need to know what Alice's choice of measurement was before Alice makes it public. Since he has no access to Alice's laboratory, doing so is impossible, even if using entangled pairs, due to nonsignaling. Alice's side of things seem much more promising. For Alice to cheat, she has to be able to produce statistics that follow either distribution (5) or distribution (6) upon request. Since the only thing she can do is perform measurements on the incoming particles, she needs to find the measurement that best allows her to mimic *both* those distributions, eventually after some postprocessing.

Note, of course, that, in an ideal setting, Alice could simply store the incoming particles in a quantum memory and measure them only when announcing her commitment. However, stable long-term quantum memories are still a long way from becoming an everyday item, and we thus assume that Alice is forced to perform her measurements on the incoming particles as soon as they arrive. For a more complete discussion, including the case of noisy memories, please refer to [3].

A. The two-state protocol

The first attempt at creating a cheating strategy is, naturally, to try and guess the state of every photon sent by Bob. Since there is no way of perfectly and systematically distinguishing nonorthogonal states, the best Alice can do, provided she is constrained to performing single-qubit measurements only, is to measure an observable "in-between" $\hat{C}_0$ and $\hat{C}_1$. Finding the basis that best discriminates states $|0\rangle$ and $|+\rangle$ is a common and well-studied problem in quantum mechanics. Such a basis is called a Breidbart basis. Specifically, let $|\tilde{0}\rangle$ ($|\tilde{1}\rangle$) denote the vector $|0\rangle$ ($|+\rangle$) rotated by $-\pi/8$ ($\pi/8$), so as to form the aforementioned Breidbart basis, and define a new observable, the *cheating observable* $\hat{C}_{\rm ch}$, by

$$\hat{C}_{\rm ch} = 0 \cdot |\tilde{0}\rangle\langle\tilde{0}| + 1 \cdot |\tilde{1}\rangle\langle\tilde{1}|, \tag{12}$$

where

$$\begin{aligned} |\tilde{0}\rangle &= \cos\frac{\pi}{8}|0\rangle - \sin\frac{\pi}{8}|1\rangle, \\ |\tilde{1}\rangle &= \sin\frac{\pi}{8}|0\rangle + \cos\frac{\pi}{8}|1\rangle. \end{aligned} \tag{13}$$

Armed with this observable, Alice has a certain probability $\mathfrak{p}_{\rm ch}(a|b)$ of obtaining the result a whenever state $|b\rangle$ is sent. Concretely, in an r-noise channel, we have

$$\begin{aligned} \mathfrak{p}_{\rm ch}(0|0) &= \frac{r}{2} + \frac{1}{4}(2 + \sqrt{2} - (2+\sqrt{2})r), \\ \mathfrak{p}_{\rm ch}(1|0) &= \frac{r}{2} + \frac{1}{4}(2 - \sqrt{2} - (2-\sqrt{2})r), \\ \mathfrak{p}_{\rm ch}(0|+) &= \frac{1}{4}(2 - \sqrt{2} + r\sqrt{2}), \\ \mathfrak{p}_{\rm ch}(1|+) &= \frac{1}{4}(2 + \sqrt{2} - r\sqrt{2}). \end{aligned} \tag{14}$$

Note that the Breidbart basis, optimal for distinguishing states $|0\rangle$ and $|+\rangle$, is also optimal for distinguishing the mixed states obtained as a result of the action of a depolarizing channel (due to the symmetry of white noise) as it saturates the Helstrom bound.

Once Alice has inferred the state of every photon, she can send her results back to Bob in hopes of passing his test. Such a strategy has been considered and analyzed in previous work [3]. However, she can do better. After all, Alice knows that her inference is not completely accurate, and so sending her inference on what the states sent by Bob may be is definitely not the best she can do. In fact, even if Alice were able to perfectly determine the state of Bob's particles, she would not send these results back to Bob but, rather, results that would emulate the statistics of an honest Alice [such as (2) or (5)]. She can thus proceed as follows: every time she obtains a 0 with the cheating observable, she decides either to change the result, with a certain probability $p_{0\to1}$, or to keep it, with probability $(1-p_{0\to1})$. Analogously, every time Alice obtains a 1, she decides either to change it, with probability $p_{1\to0}$, or to keep it, with probability $(1-p_{1\to0})$. After such a procedure, denoting by $p_{\mathrm{ch}}(a|b)$ the probability that Alice sends the value a back to Bob whenever he sends $|b\rangle$, we have

$$\begin{aligned} p_{\mathrm{ch}}(0|0) &= \mathfrak{p}_{\mathrm{ch}}(0|0)\cdot(1-p_{0\to1}) + \mathfrak{p}_{\mathrm{ch}}(1|0)\cdot p_{1\to0},\\ p_{\mathrm{ch}}(1|0) &= \mathfrak{p}_{\mathrm{ch}}(1|0)\cdot(1-p_{1\to0}) + \mathfrak{p}_{\mathrm{ch}}(0|0)\cdot p_{0\to1},\\ p_{\mathrm{ch}}(0|+) &= \mathfrak{p}_{\mathrm{ch}}(0|1)\cdot(1-p_{0\to1}) + \mathfrak{p}_{\mathrm{ch}}(1|0)\cdot p_{1\to0},\\ p_{\mathrm{ch}}(1|+) &= \mathfrak{p}_{\mathrm{ch}}(1|1)\cdot(1-p_{1\to0}) + \mathfrak{p}_{\mathrm{ch}}(0|1)\cdot p_{0\to1}. \end{aligned} \tag{15}$$

The next step is rather straightforward: determine the values of $p_{0\to1}$ and $p_{1\to0}$ that most increase Alice's chances of successfully cheating. Naturally, this depends on Bob's criterion and on whether Alice is trying to trick Bob into believing she committed to 0 or trying to trick him into believing she committed to 1. Without loss of generality, assume that Alice is trying to forge a commitment to 0 and that Bob is using the goodness-of-fit test described in Sec. II, that is, he accepts Alice's commitment to 0 if the number of 0's revealed by Alice falls within the interval $[\mu_0-3\sigma_0,\mu_0+3\sigma_0]$ whenever $|0\rangle$ was sent [see (9)] and the number of 1's within the interval $[\mu_+-3\sigma_+,\mu_++3\sigma_+]$ whenever $|+\rangle$ was sent [see (11)]. The probability that a cheating Alice passes this test is then given by

$$\begin{aligned} p_{\mathrm{ch}} = &\left(\sum_{k=\mu_0-3\sigma_0}^{\mu_0+3\sigma_0}\binom{N_0}{k}p_{\mathrm{ch}}^{k}(0|0)\cdot p_{\mathrm{ch}}^{N_0-k}(1|0)\right)\\ &\times\left(\sum_{k=\mu_+-3\sigma_+}^{\mu_++3\sigma_+}\binom{N_+}{k}p_{\mathrm{ch}}^{k}(0|+)\cdot p_{\mathrm{ch}}^{N_1-k}(1|+)\right), \end{aligned} \tag{16}$$

where $N_0=N_+=N$ are the number of $|0\rangle$ states and the number of $|+\rangle$ states sent by Bob, respectively.

Although finding an expression for the maximum of such a function is nontrivial, it is quite simple to do so numerically. In Fig. 2 we present a few plots of p_{ch} for different values of the noise parameter r, and for $M=N_0+N_+=2N=100$.

As can be seen, to obtain an optimal cheating strategy in the case of a noiseless channel, a cheating Alice should always choose $p_{0\to1}$ to be 0. For higher values of noise Alice should choose $p_{0\to1}$ to be slightly higher than 0. For $M=100$ the threshold value for r, for which Alice should choose $p_{0\to1}>0$, is above 0.52, and it decreases as M grows. As for $p_{1\to0}$, the optimal value lies in the interval (0.4,0.5) but depends on the noise parameter r. Given the optimal choice of the parameters $p_{0\to1}$ and $p_{1\to0}$, the maximal probability of cheating, $p_{\mathrm{ch}}^{\max}$,

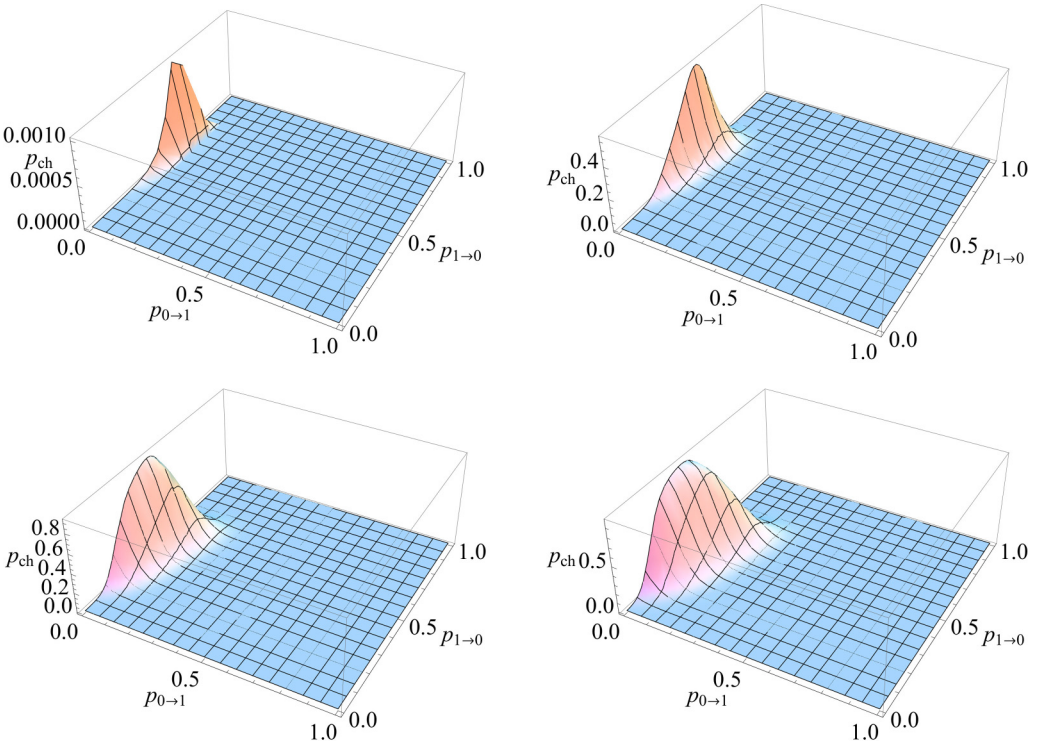


FIG. 2. Probability p_{ch} of a cheating Alice with $M=N_0+N_+=2N=100$ successful measurements passing Bob's test of commitment to 0, for different values of the noise parameter r: top left, 0; top right, 0.08; bottom left, 0.16; and bottom right, 0.24.

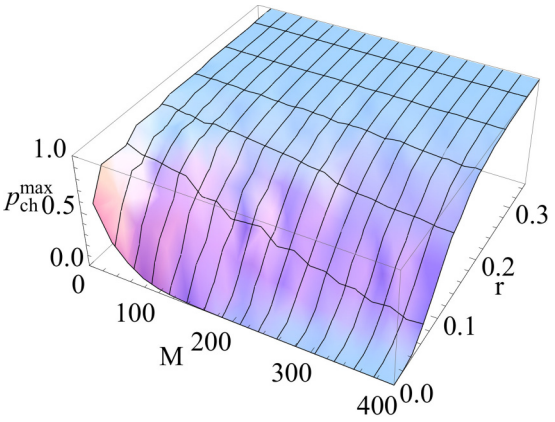


FIG. 3. Probability p_{ch}^{max} of a cheating Alice, as a function of the number $M = N_0 + N_+ = 2N$ of successful measurements, passing Bob's test of commitment to 0 and noise parameter r.

is presented in Fig. 3, as a function of $M = N_0 + N_+ = 2N$ and r.

B. The four-state protocol

It is quite interesting to compare the two-state cheating strategy above with the analogous cheating strategy for the four-state variant of the protocol. This variant, initially proposed in [4], goes as follows:

(1) Alice and Bob agree on the two observables

$$\hat{C}_0 = 0 \cdot |0\rangle\langle 0| + 1 \cdot |1\rangle\langle 1|,$$
$$\hat{C}_1 = 0 \cdot |-\rangle\langle -| + 1 \cdot |+\rangle\langle +|. \tag{17}$$

(2) Bob randomly chooses one of the four states, $|0\rangle$, $|1\rangle$, $|+\rangle$, or $|-\rangle$, and then prepares and sends a particle in that state to Alice. He repeats this step a number M of times of his choice.

(3) Alice either measures $\hat{C}_0$, for committing to 0, on all the particles she receives or measures $\hat{C}_1$, for committing to 1, on all the particles she receives.

(4) Whenever Alice wants to reveal her commitment, she sends her commitment choice along with the results of her measurements back to Bob.

(5) Bob either accepts or discards Alice's commitment, based on some statistical criterion of his choosing.

Instead of two tests, Alice now has to pass four distinct tests: one for whenever each different state was sent. The four probability distributions are now given by Eq. (5), Eq. (6), and two additional but analogous ones for whenever states $|1\rangle$ and $|-\rangle$ are sent. Again, in Fig. 4 we plot the probability that a cheating Alice, measuring all photons in the Breidbart basis (the same as for the two-state protocol), passes all four tests of commitment to 0, as a function of the probability of changing her mind about 0's ($p_{0\to1}$) and about 1's ($p_{1\to0}$), for $N = N_{0/1} = N_{+/-} = 50$.

As can be seen, it is always advantageous for a cheating Alice to change her mind about both 0's and 1's. Finally, in Fig. 5 we show the maximal probability of cheating, p_{ch}^{max}, as a function of $M = N_0 + N_1 + N_+ + N_- = 4N$ and r. Compared with the two-state case, it is clear that in terms of the resources (the total number of measurement outcomes, $M = 2N$ for two-state and $M = 4N'$ for four-state protocols, with $N' = N/2$), the four-state protocol is advantageous.

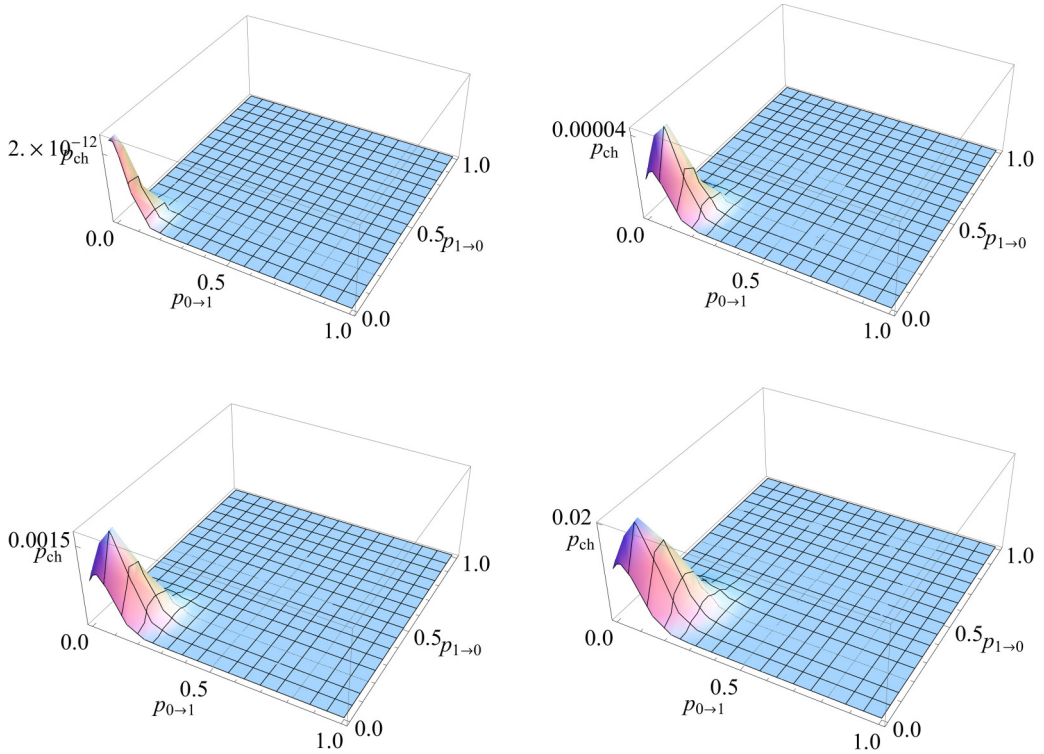


FIG. 4. Probability of a cheating Alice with $M = N_0 + N_1 + N_+ + N_- = 200$ successful measurements passing Bob's test of commitment to 0, for different values of the noise parameter r: top left, 0; top right, 0.08; bottom left, 0.16; and bottom right, 0.24.

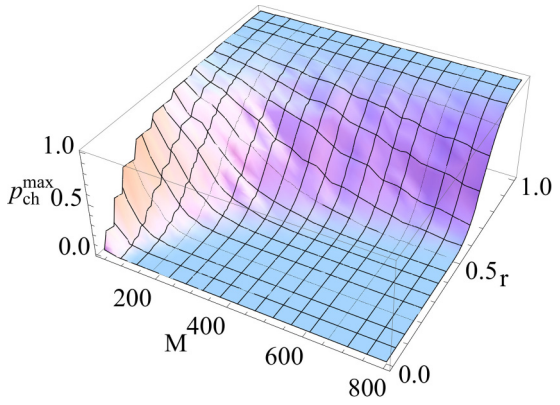


FIG. 5. Probability $p_{\text{ch}}^{\max}$ of a cheating Alice, as a function of the number $M = 4N$ of successful measurements, passing Bob's test of commitment to 0 and noise parameter r.

IV. FAKING-DISTANCE ATTACK

In any experimental implementation of the above protocols, there will be inevitable losses during transmission, just as there is inevitable noise. These losses, just like noise, need to be taken into account by Bob when receiving the final data from Alice. Indeed, if Alice is close by, then Bob expects her to receive most of the emitted particles, say, photons, and thus expects a measurement result for almost each photon sent. On the other hand, if Alice is far away, then several photons will be lost during transmission, and thus Bob will be expecting only a small number of measurement results compared to the number of photons sent. Curiously enough, this mundane phenomenon is advantageous for a cheating Alice.

Suppose Alice is much closer to Bob than she claimed to be, meaning Bob is expecting that only a fraction f of the particles sent will end up producing a measurement result. In reality, Alice is able to generate measurement results for (almost) all emitted photons. So if Alice has a way to produce statistical data in perfect agreement with those of two honest parties committing to different values for a fraction f of all particles, she can trick Bob into thinking her forged statistics are the legitimate statistics obtained by an Alice far away, who had access to only f times the particles sent.

Alice can achieve this very easily by simply measuring $\hat{C}_0$ on half of the received particles and $\hat{C}_1$ on the other half. When asked to give proof of her commitment, Alice simply sends the half of the measurement results that best suits her, thus emulating an honest Alice who would have had access to only 50% of the particles sent. This strategy is stronger than the equally common technique of unambiguous state discrimination, where Alice can learn the true state of around 29% of the particles she receives.

A cheating Alice, making her measurements right outside Bob's laboratory, while pretending to be far away, can thus cheat if Bob is expecting that only 50% or less of the particles sent end up producing a measurement result (by an honest Alice). There is thus a maximal distance up to which the protocol is secure: the distance beyond which an honest Alice is able to measure only 50% or less of the particles sent. To estimate this distance, let N_P be the total number of pulses produced by Bob. In each pulse, a photon is emitted following a certain probability distribution with mean μ. The total number of emitted photons is then $N_E = \mu N_P$. The number of photons received by an honest Alice is thus $N_R = 10^{-\alpha L/10} N_E$, where α is the attenuation coefficient and L is the length of the optical fiber. Finally, due to a finite detector efficiency η, the number of measurement outcomes of an honest Alice is $N_M = \eta N_R = \eta\mu 10^{-\alpha L/10} N_P$, while for a cheating Alice it is $N_{cM} = \eta N_E = \eta\mu N_P$. A cheating Alice, performing measurements right outside Bob's laboratory ($L = 0$), produces statistics that are compatible with her future choice of commitment on 50% of the measurement outcomes. Note that we are assuming that a cheating Alice has the same type of detector as an honest Alice, since we are dealing with a practical case scenario. Thus, the maximal distance $L_{\max}$ for which the protocol is still secure is given by $N_M = N_{cM}/2$, leading to

$$L_{\max} = \frac{10}{\alpha} \log_{10}(2). \tag{18}$$

For a typical value of $\alpha = 0.2$ (as used in, for instance, [2]), this amounts to a maximal safe distance of 15 km.

This attack shares quite some similarity to the famous photon-number splitting attack on the BB84 protocol, which can be patched using decoy states, as suggested by Hwang in [16]. Unfortunately, this remedy is not applicable in our case, since the evildoer (Eve) and the honest agent (Alice) are one and the same. We are thus left with only a maximum safe distance.

In the case of real-life noisy scenario, it is natural to assume that the "distant" Alice obtains imperfect results due to the distant noise factor r_d, while the cheating "nearby" Alice has noise given by $r_n < r_d$. A cheating Alice therefore has better statistics than an honest one and can use her "good" results on 50% of the cases to infer, say, $(50 + \delta)$% of the "lesser quality" results obtained by an honest Alice: by randomly choosing δ% of the "missing" results, a cheating Alice spoils the overall quality of her statistics, to match those of an honest Alice, which additionally decreases $L_{\max}$. In other words, we have

$$N_M = \frac{N_{cM}}{2} + \Delta N, \tag{19}$$

where $\Delta N = \delta \cdot N$ (consequently, $\delta = 1/2 - 10^{-\alpha L/10}$). The statistics $p_{c;r_d}(*|*)$, with $c \in \{0,1\}$, of an honest Alice are given by Eqs. (5) and (6), with $r = r_d$. The statistics of a cheating Alice are given by a weighted sum of the "good" results, measured on $N_{cM}/2$ photons, and random guesses, presented for ΔN photons:

$$p^{ch}_{c;r_n}(*|*) = \frac{\frac{N_{cM}}{2}}{\frac{N_{cM}}{2} + \Delta N} p_{c;r_n}(*|*) + \frac{\Delta N}{\frac{N_{cM}}{2} + \Delta N} \frac{1}{2}. \tag{20}$$

Equating, say, $p^{\text{ch}}_{0;r_n}(0|0)$ with $p_{0;r_d}(0|0)$ and using (19), one obtains the expression for the maximum safe distance:

$$L_{\max} = \frac{10}{\alpha}\left[\log_{10}(2) + \log_{10}\left(1 - \frac{r_d - r_n}{1 - r_d}\right)\right]. \tag{21}$$

Note that expression (21) is defined only for $(r_d - r_n)/(1 - r_d) < 1$. If $r_d \geqslant (r_n + 1)/2$, there exist no $L_{\max}$, i.e., a dishonest Alice can always fake her results successfully, which comes as no surprise: for such high noise, $r_d > 1/2$, the two

commitments become indistinguishable and the protocol is not viable. The above analysis is applicable to both two- and four-state protocols.

V. MULTIPHOTON SOURCES

There is still one detail that we have discarded up until now but that plagues any implementation that may currently be done: the nonexistence of perfect single-photon sources. Indeed, up until this point we have never considered the possibility that Alice may find with certainty in which state certain particles are sent by Bob. It turns out, however, that any device used to prepare photons in a specific state has a non-negligible probability of creating a pair (or even triple, quadruple, etc.) of photons in the same state instead. In a realistic scenario, one must consider a photon source whose number of photons emitted per pulse follows a Poisson distribution with a certain parameter μ (average number of photons per pulse), characteristic of the source. If we denote by $P(n,\mu)$ the probability of obtaining n photons in a given pulse, we have

$$P(n,\mu) = \frac{\mu^n}{n!}e^{-\mu}. \tag{22}$$

Alas, as soon as Bob sends more than one photon in a single pulse, which happens with probability $(1 - e^{-\mu} - \mu e^{-\mu})$, then Alice may split up the photons and measure the two observables, $\hat{C}_0$ and $\hat{C}_1$, on two of them separately, thus learning the photon's true state (provided Bob follows the protocol) and producing statistics compatible with either a commitment to 0 or a commitment to 1 upon request. If Alice receives only one photon, she reverts to the strategy described in Sec. III. We must then once again reevaluate the sets of cheating probabilities in order to obtain $p_{\mathrm{CH}}(i|j)$: the probability that a cheating Alice, using every single tool at her disposal, will send the value i back to Bob, every time he sends state $|j\rangle$. When trying to forge a commitment to 0, the corresponding set of conditional probabilities for a two-state case is (the additional conditional probabilities for a four-state case are obtained analogously)

$$\begin{aligned}
p_{\mathrm{CH}}(0|0) &= \frac{1}{1-e^{-\mu}}[\mu e^{-\mu} p_{\mathrm{ch}}(0|0) \\
&\quad + (1 - e^{-\mu} - \mu e^{-\mu}) p_0(0|0)], \\
p_{\mathrm{CH}}(1|0) &= \frac{1}{1-e^{-\mu}}[\mu e^{-\mu} p_{\mathrm{ch}}(1|0) \\
&\quad + (1 - e^{-\mu} - \mu e^{-\mu}) p_0(1|0)], \\
p_{\mathrm{CH}}(0|+) &= \frac{1}{1-e^{-\mu}}[\mu e^{-\mu} p_{\mathrm{ch}}(0|+) \\
&\quad + (1 - e^{-\mu} - \mu e^{-\mu}) p_0(0|+)], \\
p_{\mathrm{CH}}(1|+) &= \frac{1}{1-e^{-\mu}}[\mu e^{-\mu} p_{\mathrm{ch}}(1|+) \\
&\quad + (1 - e^{-\mu} - \mu e^{-\mu}) p_0(1|+)],
\end{aligned} \tag{23}$$

and analogously for the case of a commitment to 1, by replacing $p_0(*|*)$ with $p_1(*|*)$.

In more detail, take, for instance, the very first value: $p_{\mathrm{CH}}(0|0)$. This value has meaning only if Bob actually sent something, which happens with probability $(1 - e^{-\mu})$, hence the factor $\frac{1}{1-e^{-\mu}}$. Given the case, two mutually incompatible results may occur. Either Bob sends a single particle, which happens with probability $\mu e^{-\mu}$, and Alice measures it using her single-photon strategy from Sec. III [hence $p_{\mathrm{ch}}(0|0)$] or Bob sends two or more particles, which happens with probability $(1 - e^{-\mu} - \mu e^{-\mu})$, and Alice perfectly emulates the probability distribution of an honest party [hence $p_0(0|0)$]. All other expressions follow an analogous reasoning.

Note that that the conditional probabilities $p_{\mathrm{ch}}(*|*)$, corresponding to the strategy in Sec. III, are those which maximize the overall cheating probability $p_{\mathrm{ch}}^{\max}$ with respect to the case of the multiphoton sources. Thus, the optimal values of $p_{0\to1}$ and $p_{1\to0}$ are, in general, different for the strategies that maximize the cheating probability for the case of the single-photon source and the case of the multiphoton source. In Tables I and II, we compare the mentioned optimal values for both two- and four-state protocols, respectively, for $M = 100$, 200, 300, and 400, with $r = 0.1$ and, for the multiphoton case, $\mu = 0.2$ (the value used in the experimental realization of a two-state protocol [2]).

The cheating strategy considered here is beyond the capabilities of technology, both at present and in any foreseeable future. To perform this kind of attack, a cheating Alice needs far more than just a nondemolition measurement. She needs a device capable of detecting whether or not the photon source emitted a single photon and needs to change her measurement apparatus accordingly. If a single photon was sent, she measures it using the optimal cheating observable. Otherwise, she splits two of the incoming photons into two different measurement apparatuses (measuring $\hat{C}_0$ and $\hat{C}_1$, respectively). Such technology is far beyond our current capabilities. More realistic is to perform the beam-splitter strategy for all of the pulses. This way, a cheating Alice would on average measure $\hat{C}_0$ on half of the single-photon pulses

TABLE I. The two-state protocol's optimal values $(p_{0\to1},\ p_{1\to0})$ for single-photon (column 2) and multiphoton (column 3) sources, for $M = 100$, 200, 300, and 400, with $r = 0.1$ and, for the multiphoton case, $\mu = 0.2$.

M	Single photon	Multiphoton
100	(0.0, 0.489663)	(0.0, 0.492572)
200	(0.0, 0.490563)	(0.0, 0.494314)
300	(0.0, 0.479936)	(0.0, 0.483053)
400	(0.0, 0.47544)	(0.0, 0.478355)

TABLE II. The four-state protocol's optimal values $(p_{0\to1},\ p_{1\to0})$ for single-photon (column 2) and multiphoton (column 3) sources, for $M = 100$, 200, 300, and 400, with $r = 0.1$ and, for the multiphoton case, $\mu = 0.2$.

M	Single photon	Multiphoton
100	(0.0658591, 0.0658591)	(0.0470355, 0.0470355)
200	(0.0793028, 0.0793028)	(0.0592754, 0.0592754)
300	(0.0939039, 0.0939039)	(0.0743531, 0.0743531)
400	(0.101166, 0.101166)	(0.0818707, 0.0818707)

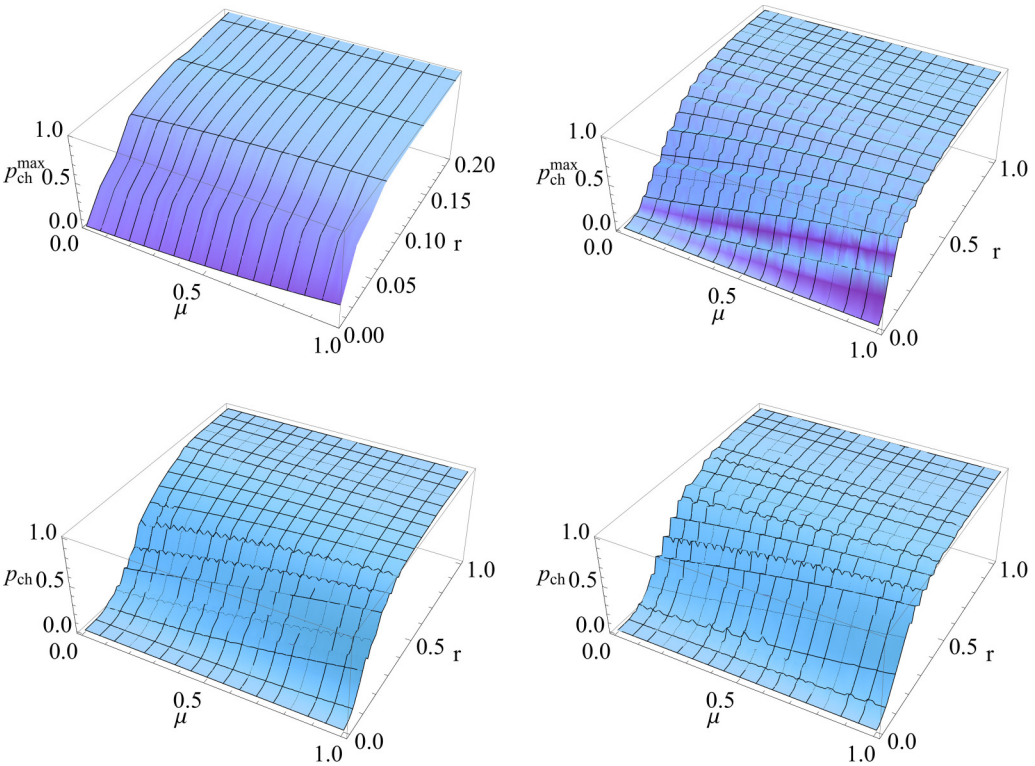


FIG. 6. Probabilities of a cheating Alice with $M = 100$ successful measurements passing Bob's test of commitment to 0, as functions of the average number of photons per pulse $\mu \in \{0,1\}$ and the noise parameter r, for the two-state protocol with the less realistic, (23), strategy (top left); the two-state protocol with the more realistic, (24), strategy (bottom left); the four-state protocol with the less realistic, (23), strategy (top right); and the four-state protocol with the more realistic, (24), strategy (bottom right).

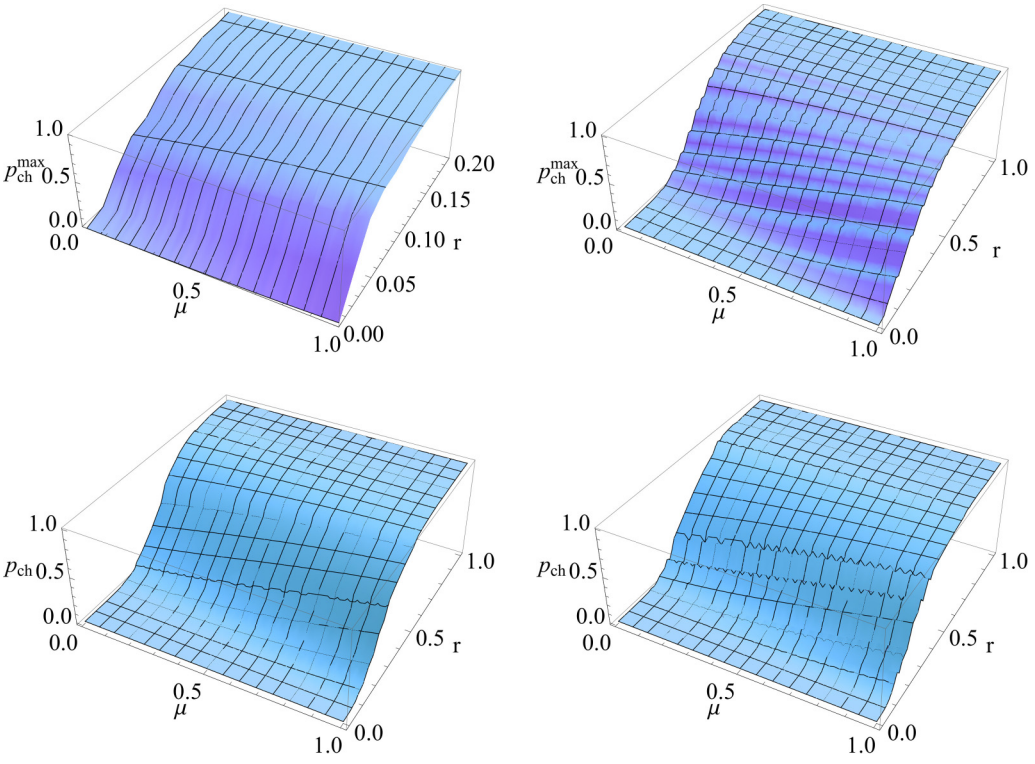


FIG. 7. Probabilities of a cheating Alice with $M = 200$ successful measurements passing Bob's test of commitment to 0, as functions of the average number of photons per pulse $\mu \in \{0,1\}$ and the noise parameter r, for the two-state protocol with the less realistic, (23), strategy (top left); the two-state protocol with the more realistic, (24), strategy (bottom left); the four-state protocol with the less realistic, (23), strategy (top right); and the four-state protocol with the more realistic, (24), strategy (bottom right).

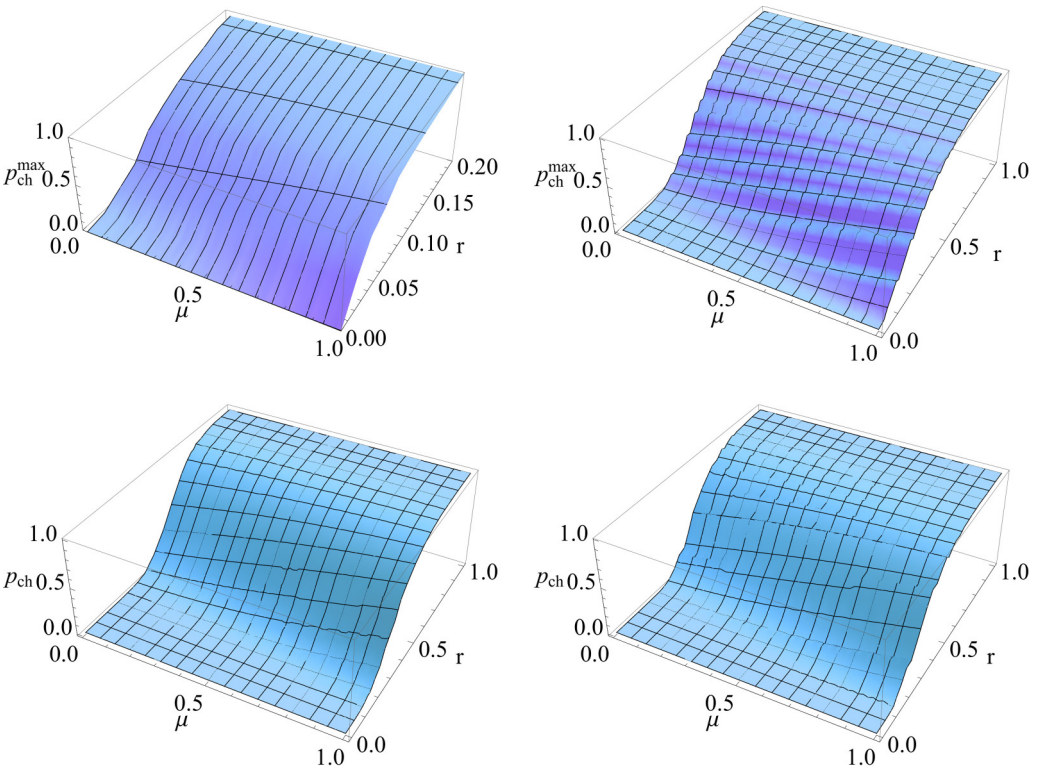


FIG. 8. Probabilities of a cheating Alice with $M = 300$ successful measurements passing Bob's test of commitment to 0, as functions of the average number of photons per pulse $\mu \in \{0,1\}$ and the noise parameter r, for the two-state protocol with the less realistic, (23), strategy (top left); the two-state protocol with the more realistic, (24), strategy (bottom left); the four-state protocol with the less realistic, (23), strategy (top right); and the four-state protocol with the more realistic, (24), strategy (bottom right).

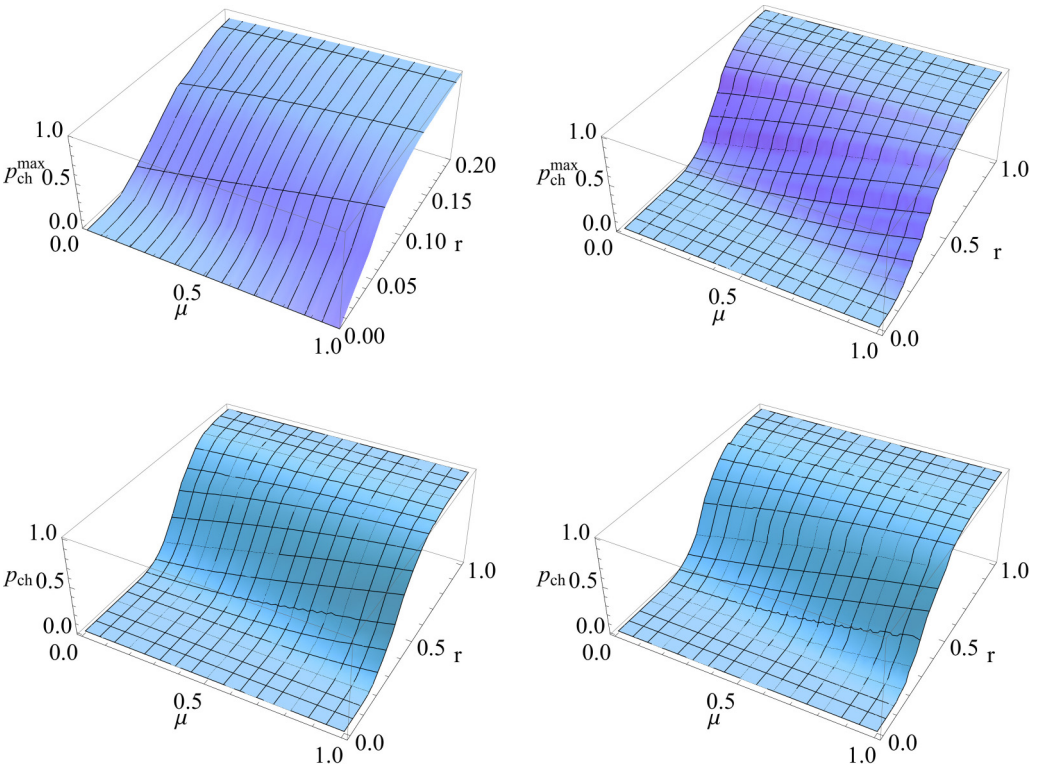


FIG. 9. Probabilities of a cheating Alice with $M = 400$ successful measurements passing Bob's test of commitment to 0, as functions of the average number of photons per pulse $\mu \in \{0,1\}$ and the noise parameter r, for the two-state protocol with the less realistic, (23), strategy (top left); the two-state protocol with the more realistic, (24), strategy (bottom left); the four-state protocol with the less realistic, (23), strategy (top right); and the four-state protocol with the more realistic, (24), strategy (bottom right).

$(\frac{1}{2}\mu e^{-\mu} N_P)$ and $\hat{C}_1$ on the other half, while for the rest of the multiphoton pulses she will have the results of both the observables. Thus, for photons on which the wrong observable was measured, a cheating Alice has to add random results. When committing to 0, her statistics for a two-state case (the additional conditional probabilities for a four-state case are obtained analogously) are given by

$$p_{BS}(*|*) = \left(1 - \frac{1}{2}\frac{\mu e^{-\mu}}{1-e^{-\mu}}\right) p_0(*|*) + \left(\frac{1}{2}\frac{\mu e^{-\mu}}{1-e^{-\mu}}\right) \cdot \frac{1}{2}, \tag{24}$$

and analogously for a commitment to 1, by replacing $p_0(*|*)$ with $p_1(*|*)$.

Naturally, the higher μ is, the closer the above two distributions, given by (23) and (24), are to those obtained by an honest Alice committing to 0 or committing to 1, respectively. Thus, we consider only weak sources with a low average photon number $\mu \in \{0,1\}$ (indeed, in a recent experimental realization of a two-state protocol [2], the value $\mu \approx 0.2$ was used). In Figs. 6, 7, 8, and 9 are the plots of the probabilities of cheating for both types of strategies [less realistic, (23), and more realistic, (24)] and both versions of the protocol (two- and four-state), as functions of the average number of photons per pulse μ and the noise parameter r, for $M = 100, 200, 300,$ and 400, respectively.

VI. CONCLUSIONS

We have analyzed the cheating strategies for the practical two-state and four-state bit-commitment protocols presented in [3,4], respectively. We show that by introducing the "postprocessing" of the raw experimental data, one can improve the naïve cheating strategy based on the measurement in the Breidbart basis, originally studied in [3]. We have studied the two cases of perfect single-photon and realistic multiphoton sources, showing that the postprocessing optimal parameters $p_{0\to1}$ and $p_{1\to0}$ are in general different for the two cases. We have also analyzed the more realistic multiphoton-source beam-splitter strategy, which does not rely on (single-photon) nondemolition measurements. In all of the mentioned cases, the four-state protocol shows a clear advantage over the two-state version, with respect to the resources needed to achieve the same level of security. Finally, we have analyzed an alternative strategy, the faking-distance attack, based on the forged geographical location of the committing party (Alice), showing that for today's typical equipment the protocol is secure up to a distance of 15 km between the parties, for both two- and four-state protocols. Our approach is, following the analysis presented in [3], straightforward to apply in the cases of noisy and/or bounded memories.

ACKNOWLEDGMENTS

The authors acknowledge financial support by the bilateral scientific co-operation between Portugal and Serbia through the project "Noise and measurement errors in multi-party quantum security protocols", No. 451-03-01765/2014-09/04 supported by the Foundation for Science and Technology (FCT), Portugal, and the Ministry of Education, Science and Technological Development of the Republic of Serbia. R.L. thanks the FCT for support through Ph.D. Grant No. SFRH/BD/79571/2011. N.P. acknowledges the support of the Security and Quantum Information Group (SQIG) and IT project QbigD funded by FCT Grants No. PEst-OE/EEI/LA0008/2013 and No. UID/EEA/50008/2013. D.A., D.P., and S.P. acknowledge the support of Ministry of Education, Science and Technological Development of the Republic of Serbia Contracts No. 171006, No. 171017, No. 171020, No. 171038, and No. 45016.

[1] Á. J. Almeida, R. Loura, N. Paunković, N. A. Silva, N. J. Muga, P. Mateus, P. S. André, and A. N. Pinto, A brief review on quantum bit commitment, in *Proceedings, SPIE* (SPIE, Bellingham, WA, 2014), pp. 92861C–92861C-8.

[2] Á. J. Almeida, A. D. Stojanovic, N. Paunković, R. Loura, N. J. Muga, N. A. Silva, P. Mateus, P. S. André, and A. N. Pinto, Implementation of a two-state quantum bit commitment protocol in optical fibers, J. Opt. **18**, 015202 (2016).

[3] R. Loura, Á. J. Almeida, P. S. André, A. N. Pinto, P. Mateus, and N. Paunković, Noise and measurement errors in a practical two-state quantum bit commitment protocol, Phys. Rev. A **89**, 052336 (2014).

[4] A. Danan and L. Vaidman, Practical quantum bit commitment protocol, Quant. Info. Proc. **11**, 769 (2012).

[5] R. Koenig, S. Wehner, and J. Wullschleger, Unconditional security from noisy quantum storage, IEEE Trans. Info. Theory **58**, 1962 (2012).

[6] N. Ng, S. Joshi, C. Ming, C. Kurtsiefer, and S. Wehner, Experimental implementation of bit commitment in the noisy-storage model, Nat. Commun. **3**, 1326 (2012).

[7] C. Schaffner, B. M. Terhal, and S. Wehner, Robust cryptography in the noisy-quantum-storage model, Quantum Inf. Comput. **9**, 963 (2011).

[8] S. Wehner, C. Schaffner, and B. M. Terhal, Cryptography from noisy storage, Phys. Rev. Lett. **100**, 220502 (2008).

[9] H. K. Lo and H. F. Chau, Is quantum bit commitment really possible? Phys. Rev. Lett. **78**, 3410 (1997).

[10] D. Mayers, Unconditionally secure quantum bit commitment is impossible, Phys. Rev. Lett. **78**, 3414, (1997).

[11] A. Souto, P. Mateus, P. Adão, and N. Paunković, Bit-string oblivious transfer based on quantum state computational distinguishability, Phys. Rev. A **91**, 042306 (2015).

[12] M. Blum, Coin flipping by telephone a protocol for solving impossible problems, in *ACM SIGACT News - A Special Issue on Cryptography* (ACM, New York, 1983), pp. 23–27.

[13] A. Shamir, R. L. Rivest, and L. M. Adleman, Mental poker, in *The Mathematical Gardner*, edited by D. A. Klarner (Springer, New York, 1981), pp. 37–43.

[14] O. Goldreich, S. Micali, and A. Wigderson, Proofs that yield nothing but their validity or all languages in np have zero-knowledge proof systems, J. ACM **38**, 690 (1991).

[15] Y. Lindell and B. Pinkas, *An Efficient Protocol for Secure Two-Party Computation in the Presence of Malicious Adversaries* (Springer, Berlin, 2007), pp. 52–78.

[16] W.-Y. Hwang, Quantum key distribution with high loss: Toward global secure communication, Phys. Rev. Lett. **91**, 057901 (2003).

IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Implementation of a two-state quantum bit commitment protocol in optical fibers

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2016 J. Opt. 18 015202

(http://iopscience.iop.org/2040-8986/18/1/015202)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 144.122.201.150
This content was downloaded on 16/01/2016 at 23:05

Please note that terms and conditions apply.

IOP Publishing Journal of Optics
J. Opt. 18 (2016) 015202 (9pp) doi:10.1088/2040-8978/18/1/015202

Implementation of a two-state quantum bit commitment protocol in optical fibers

Á J Almeida[1,2], A D Stojanovic[3], N Paunković[4,5], R Loura[4,5], N J Muga[2,6], N A Silva[2,6], P Mateus[4,5], P S André[2,7] and A N Pinto[2,6]

[1] Department of Physics, University of Aveiro, Aveiro, Portugal
[2] Instituto de Telecomunicações, Aveiro, Portugal
[3] Departamento de Engenharia de Teleinformática, Universidade Federal do Ceará, Fortaleza, Brazil
[4] SQIG—Instituto de Telecomunicações, Technical University of Lisbon, Lisbon, Portugal
[5] Department of Mathematics, Technical University of Lisbon, Lisbon, Portugal
[6] Department of Electronics, Telecommunications and Informatics, University of Aveiro, Aveiro, Portugal
[7] Department of Electrical and Computer Engineering, Technical University of Lisbon, Lisbon, Portugal

E-mail: aalmeida@av.it.pt

Received 11 May 2015, revised 21 October 2015
Accepted for publication 2 November 2015
Published 11 December 2015



Abstract

We demonstrate experimentally the feasibility of a two-state quantum bit commitment protocol, which is both concealing and partially binding, assuming technological limitations. The security of this protocol is based on the lack of long-term stable quantum memories. We use a polarization-encoding scheme and optical fiber as a quantum channel. The measurement probability for the commitment is obtained and the optimal cheating strategy demonstrated. The average success rates for an honest player in the case where the measurements are performed using equal bases are 93.4%, when the rectilinear basis is measured, and 96.7%, when the diagonal basis is measured. The rates for the case when the measurements are performed in different bases are 52.9%, when the rectilinear basis is measured, and 55.4% when the diagonal basis is measured. The average success rates for the optimal cheating strategy are 80% and 73.8%, which are way below the success rates of an honest player. Using a strict numerical validity criterion, we show that, for these experimental values, the protocol is secure.

Keywords: quantum bit commitment, quantum cryptography, polarization, optical fiber

(Some figures may appear in colour only in the online journal)

1. Introduction

Recently, there has been a strong effort to develop quantum primitives, which are designed to change classical cryptographic protocols to their quantum generalizations. One group of these primitives are quantum-based two-party computation protocols, such as the bit commitment (BC) scheme. We demonstrate the practical feasibility of a two-state quantum-bit commitment (QBC) scheme [1] based on the B92 [2] cryptographic protocol, making use of technological limitations.

Regarding current trends in cryptography and information security, QBC is believed to have higher potential than BC [3]. Unfortunately, not even the laws of quantum mechanics allow us to realize all desirable cryptographic functionalities without further assumptions. One of the most noticeable results showed that a non-relativistic QBC cannot be unconditionally secure, due to the Einstein–Podolsky–Rosen attack [4, 5]. Since then, the scientific work on QBC has been divided into theoretical and experimental. Regarding theoretical work, most attempts have been focused on evading the no-go theorem [6]. One of the first attempts extended its proof to cover ideal quantum coin tossing [7]. More complicated examples on how to apply the no-go proof to break some quantum as well as classical BC protocols that looked promising at that time were provided in [8]. Reference [9] further studied the security bounds of QBC quantitatively. Recent efforts proved the no-go theorem with alternative methods [10, 11]. Although some results were obtained in favor of QBC, until now the no-go theorem has been well

 © 2016 IOP Publishing Ltd Printed in the UK

accepted in the quantum cryptographic community, and there has been an effort to build a secure QBC under various settings [12, 13]. These impossibility results rule out the ability to build perfect cryptographic primitives. However, it could be possible to build a QBC that is almost perfect. Aharonov first showed how to construct an imperfect QBC with cheating probabilities smaller than 0.9143 [14]. The best protocol was due to Ambainis, who constructed a QBC scheme (and a quantum coin-flipping protocol) where no player can cheat, with a probability greater than 3/4 [15]. Since all implementations suffer from noise, several proposals for secure BC under the noisy-storage model were presented [16–19], being also demonstrated experimentally [20]. The practical security of QBC assuming technological limitations was also presented in [21] (security against a quantum memoryless adversary was already pointed out in the original BB84 paper [22]). However, by joining quantum mechanics and relativity, it was demonstrated to be possible to create an unconditionally secure QBC protocol based on the exchange of quantum and classical information [13, 23], which was proved to be secure against quantum adversaries [24, 25]. This protocol was already implemented [26, 27]. An information-theoretically secure commitment scheme based on [28] and allowing long commitment times was recently implemented [29]. Even if they are unconditionally secure, relativistic protocols are also difficult to implement. Based on the work from [21], a practical two-state QBC protocol, along with noise and measurement errors calculations, was recently proposed [1].

We present an experimental implementation of the two-state QBC protocol proposed in [1]. An overview of the protocol considerations along with the noise analysis is presented, as well as its experimental implementation. Our experimental work is performed using optical fibers, and the encoding of qubits uses photon polarization. The security of the protocol is based on the lack of long-term stable quantum memories. Provided this technological limitation, the security of the protocol is based on the laws of quantum mechanics and not on mathematical hardness conjectures, which are the basis of all asymmetric classical cryptographic schemes (a generic advantage quantum cryptography has over classical, shown already by the pioneering key distribution protocols of Bennet and Brassard [2, 22]).

This paper contains five sections. In section 2, we present the protocol definition along with the associated measurement probabilities. In section 3, we show the experimental setup and discuss the results. In section 4 we analyze the protocol's security against cheating attempts. Section 5 discusses the main conclusions of this work.

2. Protocol's definition

QBC is a two-party protocol which allows Bob to know about Alice's commitment. The protocol is binding; i.e., Alice cannot change her commitment after the commitment phase, and concealing; i.e., Bob cannot learn Alice's commitment before the opening phase. In practice, QBC allows one to solve trust issues when people are not close to each other. In this section we present the protocol steps along with the underlying theory of noisy channels. First, it is important to notice that in this protocol it is Alice who makes the commitment (as in the classical BC), while Bob is initializing the protocol by sending individual photons (contrary to the classical BC, where Alice is sending classical bits).

It is emphasized that the qubit states $|0\rangle$ and $|1\rangle$ used in the protocol are not orthogonal. Indeed, we choose states such that $\langle 0|1\rangle = \cos(\pi/4)$. The states orthogonal to $|0\rangle$ and $|1\rangle$ are denoted by $|0^{\perp}\rangle$ and $|1^{\perp}\rangle$, respectively, so that $\langle 0^{\perp}|0\rangle = 0$ and $\langle 1^{\perp}|1\rangle = 0$. In that way, we defined two orthonormal bases $\mathcal{B}_0 = \{|0\rangle, |0^{\perp}\rangle\}$ and $\mathcal{B}_1 = \{|1\rangle, |1^{\perp}\rangle\}$, which define two orthogonal observables that we call $\hat{C}_0$ and $\hat{C}_1$:

$$\hat{C}_0 = 0 \cdot |0\rangle\langle 0| + 1 \cdot |0^{\perp}\rangle\langle 0^{\perp}|,$$
$$\hat{C}_1 = 1 \cdot |1\rangle\langle 1| + 0 \cdot |1^{\perp}\rangle\langle 1^{\perp}|. \quad (1)$$

For the measurement strategy, Alice is measuring one of the two observables defined in equation (1) (see a detailed discussion below). Since the angle between the two states $|0\rangle$ and $|1\rangle$, and the respective bases $\mathcal{B}_0$ and $\mathcal{B}_1$, is $\pi/4$, the probability of having a mismatch between the state sent and the measurement outcome is 1/2 [30]. We now turn our attention to factors which contribute to an increased quantum-bit error rate (QBER). We will divide them into two groups. The first group will represent the effects that change the quantum state of a photon—the noise as a consequence of the photon's interaction with the environment and misalignment of the experimental apparatus. Following [31], we call it the *optical noise* and model it by a depolarizing channel (see section IV.A; note that the visibility V from equation (34) of [31] and the depolarizing channel parameter p given by equation (2) below are related by a simple relation $V = 1 - p$). The other group of factors that contribute to the QBER are imperfect single-photon sources, photon-absorption by the environment, finite detector efficiencies, and detectors' dark counts. In contrast to the first group, we will refer to these factors as *non-optical noise*.

The protocol has four phases and runs as follows:

1. **Initialization:** Bob generates a random sequence of classical bits, encodes them in one of two polarization states, $|0\rangle$ or $|1\rangle$, and sends them to Alice. Bob then keeps the record of the states of the photons sent to Alice.
2. **Commitment:** Immediately after receiving a photon from Bob, Alice performs a measurement on it. She measures one of the two observables $\hat{C}_0$ or $\hat{C}_1$ on all photons received from Bob. The choice of the measurement observable is the following: if Alice wants to commit to 0, then she measures $\hat{C}_0$; if she wants to commit to 1, then she measures $\hat{C}_1$, on all photons.
3. **Opening:** Alice reveals her commitment. She informs Bob the observable she measured and the measurement results she obtained.
4. **Validation:** Bob performs a goodness-of-fit test to check whether Alice's measurements are statistically sound and either accepts or rejects Alice's commitment based on this test (see the paragraph below).

There are various schemes that Bob can adopt to validate Alice's commitment. We here take advantage of the simplicity of the protocol and assume Bob uses the following test, known as the binomial test: let $p_c(r|b)$, with $c, r, b \in \{0, 1\}$, denote the conditional probability that Alice obtains result r when measuring observable $\hat{C}_c$ on state $|b\rangle$. All these can be easily computed or estimated by Bob before the protocol starts. Let $n(r|b)$ be the number of r's measured by Alice whenever state $|b\rangle$ was sent, and define $q(r|b) = \dfrac{n(r|b)}{n(b)}$. Then the sets $\{q(r|0)\}_{r\in\{0,1\}}$ and $\{q(r|1)\}_{r\in\{0,1\}}$ form sets of statistical data. Without loss of generality, suppose Alice is committing to 0. Let $P(q(r|0)||p_0(0|0))$ be the probability that a binomial distribution with probability of success $p_0(0|0)$ produces the statistics $\{q(r|0)\}_{r\in\{0,1\}}$, and analogously for $P(q(r|1)||p_0(1|1))$ [1]. To accept Alice's commitment, Bob checks that $P(q(r|0)||p_0(0|0)) > \alpha$ and that $P(q(r|1)||p_0(1|1)) > \alpha$, where α is a threshold probability to be determined by Bob himself. Furthermore, for viability purposes, one must require that if Alice commits to 1, she must be unable to pass the test of committing to 0. As noted in [1], this requirement is satisfied as long as it is secure against a cheating Alice. In section 4 we analyze possible choices of the security parameter α and show that the protocol is indeed secure against a cheating Alice.

To analyze the optical part of the noise, we consider a white-noise model given by a depolarizing channel $\mathcal{E}_{\mathrm{d}}$ with a given probability p; that is:

$$\mathcal{E}_{\mathrm{d}}(\hat{\rho}) = (1 - p)\hat{\rho} + p\frac{\hat{I}}{2}, \tag{2}$$

where $\hat{\rho}$ is a general mixed state that represents the initial qubit state and $\hat{I}$ is the identity operator. It is worth considering the overall conditional probabilities obtained from each measurement. If by $p_c(r|b)$, with $c, r, b \in \{0, 1\}$, we denote a conditional probability that a result r is obtained when measuring observable $\hat{C}_c$ on state $|b\rangle$, then the conditional probabilities are given by the following expressions [1],

- If Alice measures $\hat{C}_0$:

$$\begin{aligned} p_0(0|0) &= 1 - \frac{p}{2} \\ p_0(1|0) &= \frac{p}{2} \\ p_0(0|1) &= 1/2 \\ p_0(1|1) &= 1/2. \end{aligned} \tag{3}$$

- If Alice measures $\hat{C}_1$:

$$\begin{aligned} p_1(0|0) &= 1/2 \\ p_1(1|0) &= 1/2 \\ p_1(0|1) &= \frac{p}{2} \\ p_1(1|1) &= 1 - \frac{p}{2}. \end{aligned} \tag{4}$$

Let $n_c(r|b)$, with $c, r, b \in \{0, 1\}$, be the number of results r when measuring observable $\hat{C}_c$ on state $|b\rangle$, and let $n_c(b)$ be the total number of detected photons sent in the state $|b\rangle$ when measuring $\hat{C}_c$. If Alice is honest, i.e., if she indeed measures only one out of the two commitment observables $\hat{C}_c$ on all qubits received from Bob, the statistics of her measurements will, due to the law of large numbers, approach the corresponding conditional probabilities in equations (3) and (4). Thus $\dfrac{n_0(0|0)}{n_0(0)} \approx p_0(0|0) = 1 - \dfrac{p}{2} = \mathrm{SRATE}_0^{\mathrm{opt}}$ and $\dfrac{n_0(1|0)}{n_0(0)} \approx p_0(1|0) = \dfrac{p}{2} = \mathrm{QBER}_0^{\mathrm{opt}}$, and analogously for other cases. We see that the first two expressions from equation (3) represent the optical contribution (opt) for the success (SRATE) and the QBERs, respectively, obtained when measuring observable $\hat{C}_0$ on state $|0\rangle$ (and analogously for the last two expressions of equation (4)). Therefore, if the statistics of Alice's measurements approach equation (3), Bob accepts a commitment to 0; if they approach equation (4), Bob accepts a commitment to 1; otherwise Bob aborts the protocol. Note that up to this point we discussed only the so-called optical errors, due to external noise occurring during the emission, transmission, and measurement of photons.

As mentioned, the aim of the protocol is to be both binding and concealing. It is also worth mentioning that (in spite of recent efforts [32, 33]) long-term storage of photons is not going to be feasible in the foreseeable future. Therefore, Alice is forced to perform measurements as the photons arrive: she cannot postpone her commitment until after the commitment phase. Since an ideal single-photon source is very difficult to obtain, Bob is forced to use weak pulses to prevent the well-known splitting attack; the sending pulse rate should be chosen in a way such that the emission rate of multi-photons is much smaller than the emission rate of single photons. When several photons are emitted, all photons are in the same pure state of polarization. Otherwise, Alice could split two and measure $\hat{C}_0$ on one photon and $\hat{C}_1$ on another. This way she could have results consistent with both commitments (and could thus postpone her decision). We also make the basic assumption that Bob does not have access to Alice's laboratory, and since she is revealing her commitment only during the opening phase, Bob does not get any knowledge about Alice's measurement before that phase. Therefore, since after the completion of the initialization phase there is no information exchange between the two parties until the start of the opening phase, the protocol is concealing.

Regarding Alice's measurements, one can also discuss her optimal cheating strategy for single-photon measurements. The optimal cheating observable is the observable 'between' $\hat{C}_0$ and $\hat{C}_1$; i.e., it is observable when the eigenbasis is rotated by $-\pi/8$ from $\mathcal{B}_0$ (for the analysis of optimal cheating strategies in the context of quantum cryptography, see, for example, [9, 34–36]). The cheating observable $\hat{C}_{\mathrm{ch}}$ is defined by mutually orthogonal vectors, $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$, such that the angle between $|0\rangle$ and $|\tilde{0}\rangle$, and the angle between $|1\rangle$ and $|\tilde{1}\rangle$ is $\pi/8$,

$$\hat{C}_{\mathrm{ch}} = 0 \cdot |\tilde{0}\rangle\langle\tilde{0}| + 1 \cdot |\tilde{1}\rangle\langle\tilde{1}|. \tag{5}$$

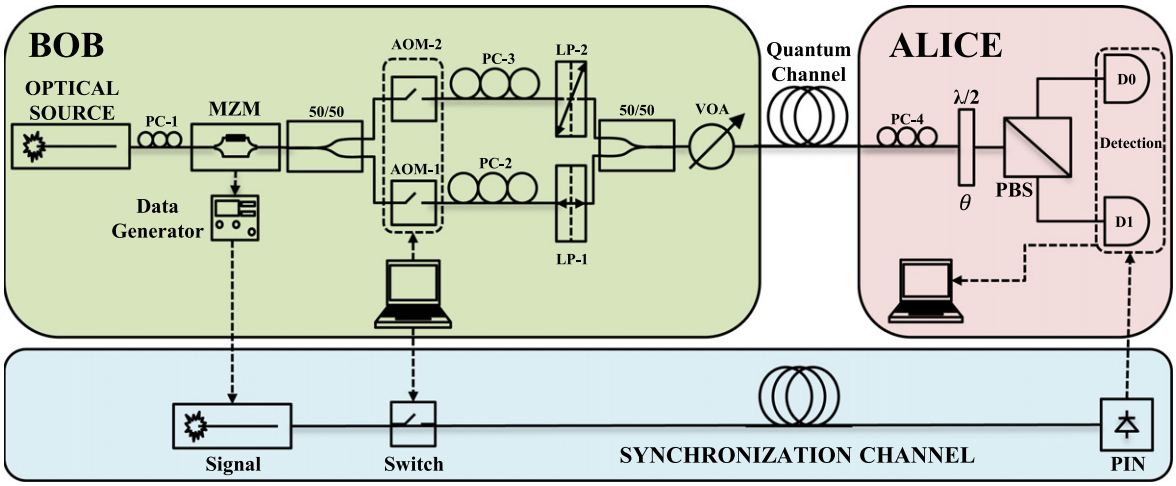


Figure 1. Experimental setup used in the proof-of-principle demonstration of the two-state QBC protocol proposed.

If Alice obtains the result 0 (corresponding to vector $|\tilde{0}\rangle$), she infers that the state sent by Bob is $|0\rangle$. When we obtain the result 1 (corresponding to state $|\tilde{1}\rangle$), we infer that the state sent by Bob is $|1\rangle$. In the ideal case, the conditional probabilities for the cheating observable are written as,

$$\begin{aligned} p_{\mathrm{ch}}(0|0) &= |\langle 0|\tilde{0}\rangle|^2 = \cos^2\left(\frac{\pi}{8}\right) \\ p_{\mathrm{ch}}(1|0) &= |\langle 0|\tilde{1}\rangle|^2 = \sin^2\left(\frac{\pi}{8}\right) \\ p_{\mathrm{ch}}(0|1) &= |\langle 1|\tilde{0}\rangle|^2 = \sin^2\left(\frac{\pi}{8}\right) \\ p_{\mathrm{ch}}(1|1) &= |\langle 1|\tilde{1}\rangle|^2 = \cos^2\left(\frac{\pi}{8}\right). \end{aligned} \tag{6}$$

In this case, the success rate is the probability to infer 0 when the state sent by Bob was $|0\rangle$, and the same for the case when the state sent by Bob was $|1\rangle$. Since the situation is symmetric, the two probabilities are the same. In terms of probabilities, $\mathrm{SRATE}_{\mathrm{ch}}^{\mathrm{opt}}$ and the $\mathrm{QBER}_{\mathrm{ch}}^{\mathrm{opt}}$ are written as:

$$\begin{aligned} \mathrm{SRATE}_{\mathrm{ch}}^{\mathrm{opt}}(0) &= \frac{n_{\mathrm{ch}}(0|0)}{n_{\mathrm{ch}}(0)} \approx p_{\mathrm{ch}}(0|0) = 0.8536 \\ \mathrm{QBER}_{\mathrm{ch}}^{\mathrm{opt}}(0) &= \frac{n_{\mathrm{ch}}(1|0)}{n_{\mathrm{ch}}(0)} \approx p_{\mathrm{ch}}(1|0) = 0.1464 \\ \mathrm{QBER}_{\mathrm{ch}}^{\mathrm{opt}}(1) &= \frac{n_{\mathrm{ch}}(0|1)}{n_{\mathrm{ch}}(0)} \approx p_{\mathrm{ch}}(0|1) = 0.1464 \\ \mathrm{SRATE}_{\mathrm{ch}}^{\mathrm{opt}}(1) &= \frac{n_{\mathrm{ch}}(1|1)}{n_{\mathrm{ch}}(0)} \approx p_{\mathrm{ch}}(1|1) = 0.8536. \end{aligned} \tag{7}$$

If the error induced by noise produces the error for an honest player (Alice) identical to the error of a cheating player, then the two strategies (honest and cheating) are indistinguishable, and the protocol cannot be performed. Thus, the honest player error rate $\mathrm{QBER}_{\mathrm{ch}}^{\mathrm{opt}}$ must be smaller than that of a cheating player, which is $\approx 15\%$.

3. Experimental implementation

In this section we present the experimental setup used to implement the two-state QBC protocol, the corresponding experimental results, and their analysis.

3.1. Experimental setup

The scheme of the experimental setup that we have used is presented in figure 1. On Bob's side, a pump at $\lambda_{\mathrm{p}} = 1550.92$ nm from a tunable laser source passes through a polarization controller (PC-1) and is externally modulated using a Mach–Zehnder modulator to produce optical pulses with a full width at half maximum of approximately 1 ns and a repetition rate of 100 kHz. The pulse goes into a 50/50 beam splitter, and two acousto-optic modulators (AOM-1 and AOM-2) will work as a switch. The PC- 2 and PC-3 are used to adjust the polarizations of the photons in the lower and upper arms, respectively. The axis of the linear polarizer LP-1 is set at 0°, preparing the $|0\rangle$ polarization state. Analogously, the axis of the linear polarizer LP-2 is set at 45°, preparing the $|1\rangle$ polarization state. Next, photons pass through a 50/50 beam coupler and are attenuated using a variable optical attenuator, obtaining an average number $\mu \approx 0.2$ of photons per pulse (such as in [31]), thus decreasing the probability of having beam-splitter attacks. Then, they reach an optical fiber (single-mode fiber with an attenuation coefficient, $\alpha \approx 0.2$ dB/km) that works as a quantum channel.

At the receiving side, Alice can choose the commitment basis by setting a half-wave plate ($\lambda/2$, with λ being the wavelength of the wave) to $2\theta = 0°$, 45°, or −22.5° and uses a polarization-beam splitter (PBS) to discriminate between 0° and 90°. The PC-4 is used to compensate random rotations of polarization in the quantum channel. Regarding Alice's commitment, if the wave plate is set at $2\theta = 0°$, it means Alice chooses $\hat{C}_0$; if it is set at $2\theta = 45°$, it means Alice chooses $\hat{C}_1$; and if it is set at $2\theta = -22.5°$ it means Alice chooses $\hat{C}_{\mathrm{ch}}$. The photons are detected using two avalanche photodiodes (D0 and D1) [37]. D0 (id200) has a dark count probability per time gate, $t_{\mathrm{g}} = 5$ ns, of $P_{\mathrm{dc}}^{0} = 3 \times 10^{-5}$ and

a quantum detection efficiency $\eta_0 \approx 7\%$. D1 (id201) has a dark count probability per time gate, $t_g = 5$ ns, of $P_{dc}^1 = 6 \times 10^{-5}$ and a quantum detection efficiency $\eta_1 \approx 9\%$. A synchronization pulse at $\lambda_s = 1547.72$ nm from a distributed feedback laser is used to trigger the detectors, where a switch allows one to send pulses with the same repetition rate of the encoded photons. These pulses are detected using a photodiode detector, which is connected to both detectors. In the synchronization channel, an optical fiber is used with the same length as the quantum channel in order to allow synchronization between Bob and Alice's measurement devices.

3.2. Experimental results

To demonstrate the feasibility of the experimental setup shown in figure 1, we have performed measurements of both $\hat{C}_0$ and $\hat{C}_1$. The rates obtained by Alice when the basis of her measurement observable $\hat{C}_c$ (with $c \in \{0, 1\}$) coincides with the basis $\mathcal{B}_b$ (with $b \in \{0, 1\}$), from which the state was prepared by Bob ('measurements in equal bases'), as well as when they are different ('measurements in different bases'), are plotted for both cases. The sequence of bits to encode was given by a pseudo-random binary sequence with 2^{17} bits. In a single run we sent the full sequence of bits and recorded the results. In figure 2 we show the measurement results obtained when Alice and Bob had a quantum channel of 16 km between them, for both $\hat{C}_0$ and $\hat{C}_1$, as a function of the run. A run means that Bob sent a pattern with the 2^{17} bits through the system, and Alice has performed the measurements.

From the experimental results obtained, we can infer several conclusions. First, we notice that the success rates for the case when the measurements are performed using equal bases are always higher than 90%, which is above the minimum theoretical security limit. The average success rate is 93.4% when $\hat{C}_0$ is measured and 96.7% when $\hat{C}_1$ is measured. Second, the rates for the case when the measurements are performed in different bases are always close to 50%, as expected. In this case, the average success rate is 52.9% when $\hat{C}_0$ is measured and 55.4% when $\hat{C}_1$ is measured. The results are according to theory, since in the ideal case, when the measurements are performed in equal bases, the success rate should be 100%. For measurements in different bases, in the ideal case the rate should be 50%, since the probability to detect a photon is the same for D0 or D1. As described in equations (3) and (4), in the realistic case of noisy channels the alignment between Bob and Alice is not perfect, which will lead to wrong detections, and consequently to a rate lower than 100%, or different from 50%, depending if we are talking about measurements in equal bases or different bases, respectively.

The SRATEs and QBERs are a consequence of both optical (depolarizing channel) and non-optical noises (imperfect single-photon sources, photon absorption by the environment, finite detector efficiency, and dark counts). Thus,

$$\mathrm{SRATE}_i = \mathrm{SRATE}_i^{\mathrm{opt}} + \mathrm{SRATE}_i^{\mathrm{non-opt}}, \tag{8a}$$

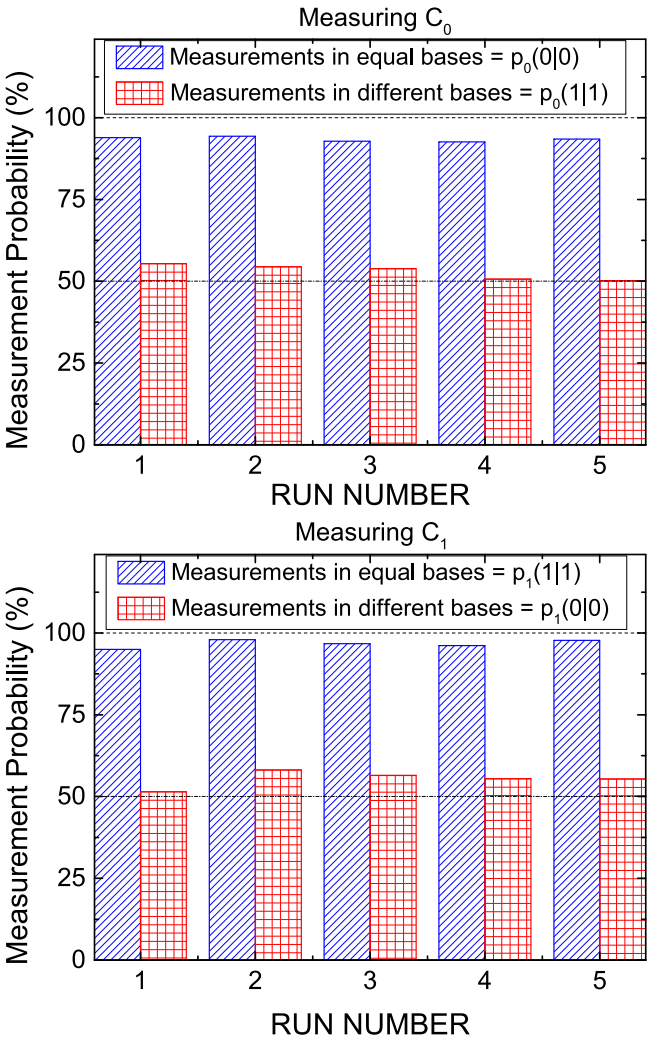


Figure 2. Experimental results of the measurement probability for five different runs obtained when Alice and Bob are interconnected with a quantum channel of 16 km. The dashed line represents the theoretical value for measurements in equal bases, and the dash-dotted line represents the theoretical value for measurement in different bases.

$$\mathrm{QBER}_i = \mathrm{QBER}_i^{\mathrm{opt}} + \mathrm{QBER}_i^{\mathrm{non-opt}}, \tag{8b}$$

with $i = 0, 1$. Theoretically, the non-optical QBER can be written as,

$$\mathrm{QBER}^{\mathrm{non-opt}} = \frac{P_{dc}^i}{\mu t_{\mathrm{link}} \eta_i}, \tag{9}$$

where μ is the average number of photons per pulse and $t_{\mathrm{link}} = 10^{(-\alpha L/10)}$, with L being the fiber length (see equations (31)–(33) in [31], page 166).

The optical noise is given by equations (3) and (4) from section 2 and plays an important role in the results. The non-optical noise does not depend on the state emitted by the source (Bob's apparatus) nor the observable measured by Alice, but it does depend on the result obtained. However, it is worth mentioning that both detectors have different detection efficiencies ($\eta_0 \approx 7\%$ for D0 and $\eta_1 \approx 9\%$ for D1),

and their dark count probabilities per time gate are also different ($P_{\rm dc}^0 = 3 \times 10^{-5}$ for D0 and $P_{\rm dc}^1 = 6 \times 10^{-5}$ for D1).

Using equations (3), (4), and (9), (8*b*) can be rewritten as,

$$\mathrm{QBER}_i = \frac{p_i}{2} + \frac{P_{\rm dc}^i}{\mu t_{\rm link}\eta_i}, \tag{10}$$

which includes both optical and non-optical contributions.

In addition to figure 2, we can plot the measurement results for the average QBER over the number of runs for several fiber lengths and use the variables p_i in equation (10) as fitting parameters, as shown in figure 3. As expected, the QBER increases with the fiber length. The difference between the QBER values in the measurements in $\hat{C}_0$ and $\hat{C}_1$ is due to several reasons: different dark count probabilities and quantum efficiencies of the two detectors and different losses in each channel.

Rewriting equation (10), we obtain

$$p_i(L) = 2\left(\mathrm{QBER}_i - \frac{P_{\rm dc}^i}{\mu t_{\rm link}\eta_i}\right). \tag{11}$$

From this equation, using the experimental results from figure 3, where each QBER_i is the average value of five runs for each fiber length, we plot the values of $p_i(L)$ in figure 4, which show the probability of the occurrence of white noise.

In order to check the optimal cheating strategy for Alice, we performed measurements of $\hat{C}_{\rm ch}$. The results are shown in figure 5. From this figure it can be seen that there is a relatively good agreement between experimental results and theoretical predictions. In figure 6 we present the experimental results of the QBER for measurements of $\hat{C}_{\rm ch}$ for several fiber lengths. We can see that the QBER is a bit higher than the theoretical value, mainly because of the difficulty in setting the half-wave plate ($\lambda/2$) at exactly $2\theta = -22.5°$ and the unavoidable optical noise. The differences between measurements of ones and zeros are also due to efficiency and dark count differences in the two detectors and different losses in each channel. In fact, the same type of error due to bad alignment exists when measuring $\hat{C}_0$ and $\hat{C}_1$, but since it is much easier to align the equipment for those two measurements, we can assume that when measuring $\hat{C}_0$ and $\hat{C}_1$, this type of error is negligible with respect to the measurement of the cheating observable. The average rates for the optimal cheating strategy were 80% when measuring zeros and 73.8% when measuring ones.

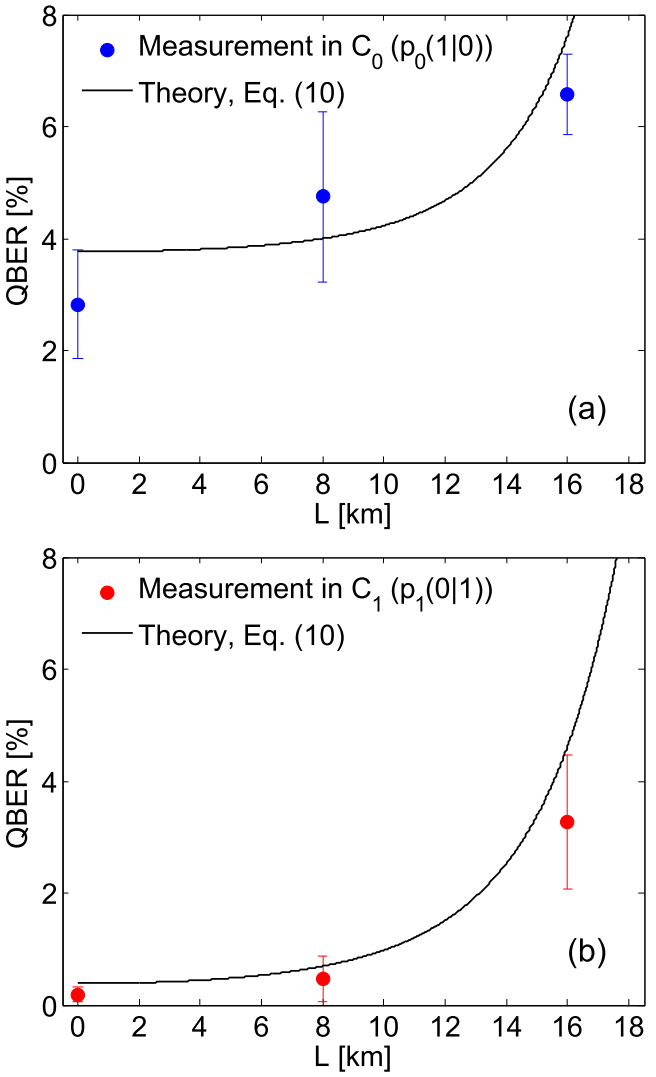


Figure 3. Average QBER as a function of the fiber length when measuring (a) $\hat{C}_0$ (QBER_0), and (b) $\hat{C}_1$ (QBER_1) along with the theoretical fit from equation (10), where $p_0 = 7.1 \times 10^{-2}$ and $p_1 = 1 \times 10^{-3}$ were used as fitting parameters. The error bars represent the standard deviation of the experimental values. The correlation coefficient between experimental and theoretical data was $r^2 = 0.77$ when $\hat{C}_0$ was measured and $r^2 = 0.99$ when $\hat{C}_1$ was measured.

4. Security against a cheating Alice

In this section we analyze the security of the protocol against the above-described cheating Alice. First, we use Alice's measurement results of observables $\hat{C}_0$ and $\hat{C}_1$ (see figure 2) to establish Bob's precise quantitative validity criterion. Then we show that the experimentally obtained results of a cheating Alice would not pass such a criterion, showing that the protocol is secure. First, we describe how Bob forms the viability criterion introduced in section 2.

The probability that an honest Alice obtains n_0 times the value 0 when measuring N_0 particles in state $|0\rangle$ is given by a binomial distribution with mean value

$$\mu_0 = N_0 p_0(0|0) \tag{12}$$

and variance

$$\sigma_0^2 = N_0 p_0(0|0)\big(1 - p_0(0|0)\big). \tag{13}$$

Since, for sufficiently large N_0, this binomial distribution behaves like a normal distribution with the same parameters, the binomial test described in section 2 reduces to the well-known $68 - 95 - 99.7$ rule of thumb for significance levels of 32%, 5%, and 0.3%, respectively. For a significance level

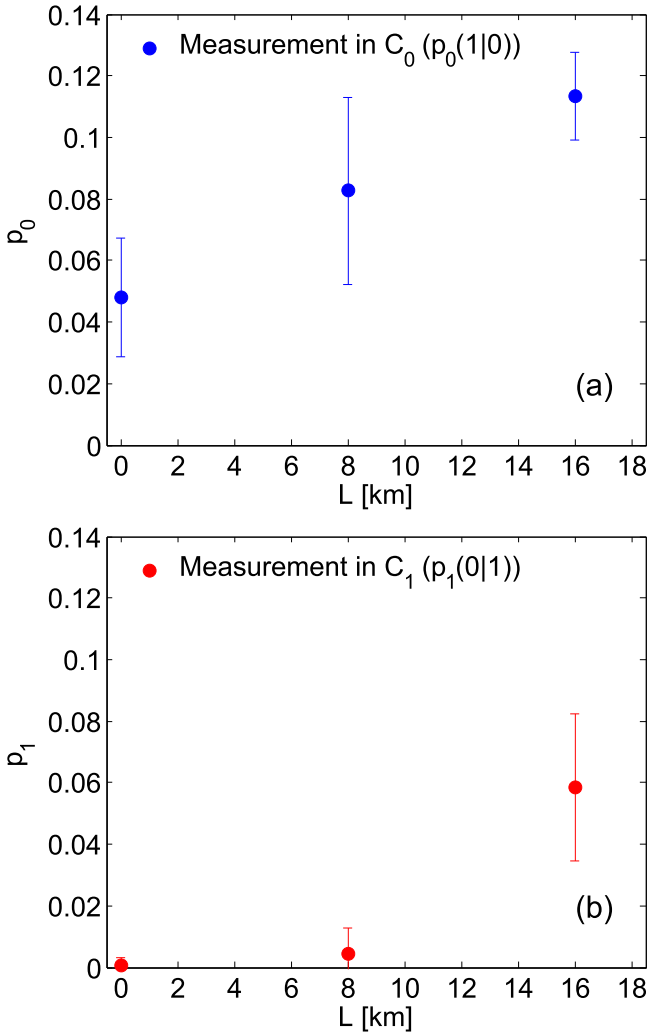


Figure 4. Calculated values for (a) p_0 and (b) p_1 as a function of L, using equation (11). The error bars represent the standard deviation of the experimental values.

of 0.3%, Bob accepts Alice's statistics for whenever $|0\rangle$ was sent if

$$n_0 \in \left[\mu_0 - 3\sigma_0,\, \mu_0 + 3\sigma_0\right], \tag{14}$$

which will henceforth be called the '3σ security criterion'. This ensures that if an honest Alice is committing to 0, then she has ≈99.7% chance of having the first part of her data set accepted by Bob as a valid commitment. In terms of the security parameter α introduced in section 2, this corresponds to setting

$$\alpha = \Pr\left[n_0 = \mu_0 - 3\sigma_0\right]. \tag{15}$$

Analogously, Bob now proceeds to test the second portion of Alice's statistics, when state $|1\rangle$ was sent, and accepts them if

$$n_1 \in \left[\mu_1 - 3\sigma_1,\, \mu_1 + 3\sigma_1\right]. \tag{16}$$

Finally, Bob validates Alice's commitment as a whole if both conditions (14) and (16) are satisfied. For the '3σ security criterion', the protocol will fail to accept an honest Alice's commitment around six times every thousand runs.

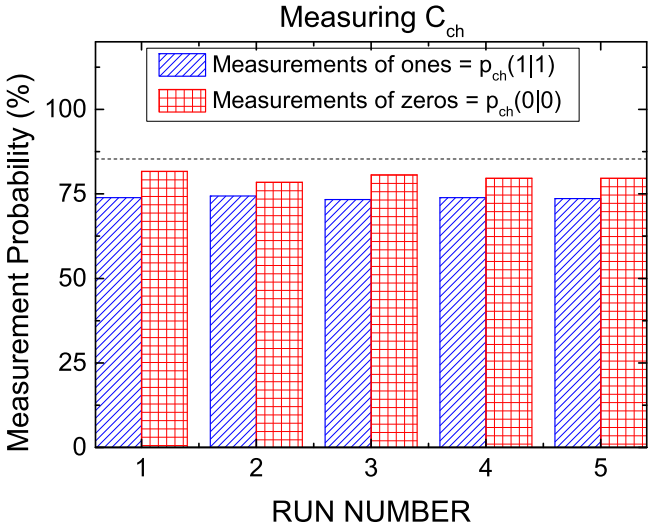


Figure 5. Experimental results of the measurement probability for five different runs obtained from the measurements in $\hat{C}_{\rm ch}$, when Alice and Bob are interconnected with a quantum channel with 16 km. The dashed line represents the theoretical value for measurements of both zeros and ones.

Concretely, figure 2 represents the measured conditional probabilities $p_0(r|b)$ when taking into account both optical and non-optical errors. The exact values are given by

$$\begin{aligned} p_0(0|0) &= 0.934 \\ p_0(1|0) &= 0.066 \\ p_0(0|1) &= 0.471 \\ p_0(1|1) &= 0.529. \end{aligned} \tag{17}$$

In this case, given that in each run of the protocol there were on average about $N_0 \approx N_1 \approx 350$ measurement outcomes, we have

$$\begin{aligned} \mu_0 &= 326.9 \\ \sigma_0 &= 4.64 \\ \mu_1 &= 185.15 \\ \sigma_1 &= 9.34. \end{aligned} \tag{18}$$

In particular, we have $[\mu_0 - 3\sigma_0,\, \mu_0 + 3\sigma_0] = [312.98,\, 340.82]$.

We now compute the probability that a cheating Alice passes the previous goodness-of-fit test. The exact values represented in figure 5 are

$$\begin{aligned} p_{\rm ch}(0|0) &= 0.8 \\ p_{\rm ch}(1|0) &= 0.2 \\ p_{\rm ch}(0|1) &= 0.262 \\ p_{\rm ch}(1|1) &= 0.738. \end{aligned} \tag{19}$$

It suffices to note that for $N_0 \approx N_1 \approx 350$ measurement outcomes obtained when measuring $\hat{C}_{\rm ch}$, a cheating Alice would obtain around $280 = N_0 p_{\rm ch}(0|0)$ times the value 0. Since $280 \notin [312.98,\, 340.82]$, the statistics from a cheating Alice would not be accepted by Bob as a valid commitment to 0. Note that even if we had chosen $[\mu_0 - 10\sigma_0,\, \mu_0 + 10\sigma_0]$ as the acceptance interval (significance level of 10^{-15}), so as

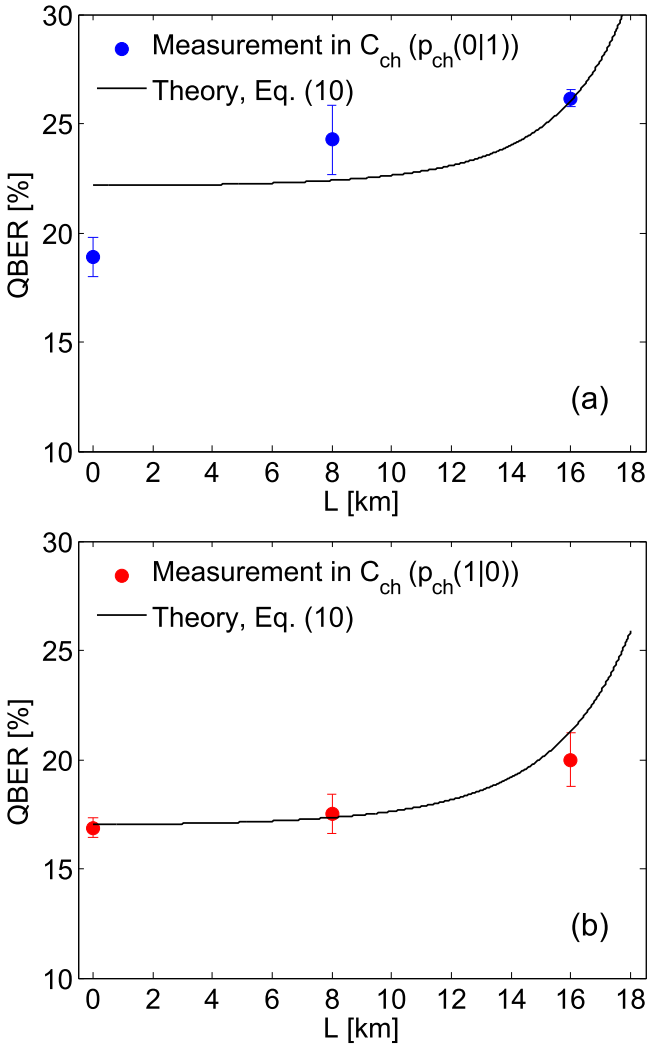


Figure 6. Average QBER as a function of the fiber length when measuring (a) ones in $\hat{C}_{\mathrm{ch}}$ and (b) zeros in $\hat{C}_{\mathrm{ch}}$, along with the theoretical fit from equation (10), where $p_0 = 0.439$ and $p_1 = 0.334$ are average values. The error bars represent the standard deviation of the experimental values. The correlation coefficient between experimental and theoretical data was $r^2 = 0.43$ when ones were measured in $\hat{C}_{\mathrm{ch}}$ and $r^2 = 0.98$ when zeros were measured in $\hat{C}_{\mathrm{ch}}$.

to increase the chance of an honest Alice being rightfully accepted, a cheating Alice would, on average, still fail to pass that test.

Note that such a low allowance for statistical error shows that, in a real-life implementation, one can use far fewer photon pulses than those used here: $2^{17} \approx 130000$. Indeed, in [1], section V, it was shown that to guarantee a 3 standard-deviation allowance, it is enough to have around $N_0 \approx N_1 \approx 50$ measurement outcomes, which corresponds to about $2^{13} \approx 8000$ photon pulses. In fact, the above optimal probability to cheat for single-qubit measurements decreases exponentially with respect to the number of photons (see, for example, theorem 4 of [38]); i.e., it is given by a *negligible function*. (A function $\varepsilon : \mathbb{N} \to \mathbb{R}$ is said to be negligible if for every positive polynomial p there exists a $n_0 \in \mathbb{N}$ such that for all $n > n_0$, $\varepsilon(n) \leqslant 1/p(n)$. See, for example, [39].)

It is interesting to note that even having perfect detectors (with perfect efficiency and no dark counts) cannot help a cheating Alice. Note that perfect detectors will only improve the results of $\hat{C}_0$ and $\hat{C}_1$ measurements used to establish Bob's numerical verification criterion, reducing them to the case of optical noise only, given by equations (3) and (4). In other words, Bob's verification criterion will be even tighter than in the case of non-ideal detectors. While perfect detectors will also improve the chances of a cheating Alice, they will not be enough to significantly compromise the protocol's security. Indeed, assuming the 'noisy' verification criterion given by equation (14), for the experimental values given by equations (17) and (18), and perfect cheating statistics given by equation (6), the protocol's security significance level is still of the order of 10^{-8}, which corresponds to a '6σ security criterion' (i.e., even a cheating Alice with exceptional access to perfect technology is successful only once in around every 100 million runs).

A general theoretical discussion of the protocol's security and viability was presented in [1], where it was shown that if the protocol is secure against a cheating Alice, then it is also viable: an honest Alice trying to commit to 1 can never pass the test of committing to 0.

5. Conclusions

We have implemented a two-state QBC protocol in optical fibers. The encoding of qubits was performed using two nonorthogonal states of polarization. We were able to implement the protocol even when photons travel through an optical fiber with several kilometers in length. We also presented results obtained for the optimal cheating strategy for Alice, showing good agreement with the theory. We noticed that the success rates for the case when the measurements were performed using equal bases were always higher than 90%. The rates for the case when the measurements were performed in different bases were always close to 50%. In this case, the average success rate was 52.9% when $\hat{C}_0$ was measured and 55.4% when $\hat{C}_1$ was measured. When measuring the observable for the optimal cheating strategy, the average success rates were 80% when measuring zeros and 73.8% when measuring ones. We presented a detailed numerical validity criterion for Bob to either accept or reject Alice's commitment and showed that, with the above-mentioned experimental rates, the protocol is secure against cheating attempts.

Finally, we note that although we presented the experimental implementation of a two-state protocol, with the necessary adjustments it can be converted to a four-state protocol. Note that in this case Bob's criterion would be split into four different conditions, similar to equations (14) and (16), thus making it harder for a cheating Alice to succeed. In general, the more states there are, the fewer photons are needed in each run of the protocol to ensure a given level of security. A rigorous theoretical analysis of the dependence of the level of security with respect to the number of photons

used can be done by following the arguments in [1], which can be straightforwardly adapted to the four-state protocol.

Acknowledgments

This work was supported in part by the FCT—Fundação para a Ciência e a Tecnologia through the PhD Grants SFRH/BD/79482/2011 and SFRH/BD/79571/2011 and the Post Doctoral Grants SFRH/BPD/77286/2011 and SFRH/BPD/96457/2013, by the FCT and the Instituto de Telecomunicações under the PEst-OE/EEI/LA0008/2011 program, project 'P-Quantum,' the PEst-OE/EEI/LA0008/2013 program, project 'CV-Quantum,' and by the Conselho de Reitores das Universidades Portuguesas (CRUP), project 'Ação Integrada E 91/12.'

The authors would like to thank Eng. João Prata from the Instituto de Telecomunicações, Aveiro, for his contribution regarding electronics.

References

[1] Loura R, Almeida Á J, André P S, Pinto A N, Mateus P and Paunković N 2014 *Phys. Rev.* A **89** 052336
[2] Bennett C H 1992 *Phys. Rev. Lett.* **68** 3121–4
[3] Zoller P *et al* 2005 *Eur. Phys. J.* D **36** 203–28
[4] Lo H K and Chau H F 1997 *Phys. Rev. Lett.* **78** 3410–13
[5] Mayers D 1997 *Phys. Rev. Lett.* **78** 3414–17
[6] Bell J S 1964 *Physics* **1** 195–200
[7] Lo H K and Chau H F 1998 *Physica* D **120** 177–87
[8] Brassard G, Crépeau C, Mayers D and Salvail L 1998 arXiv: quant-ph/9806031
[9] Spekkens R W and Rudolph T 2002 *Quantum Inf. Comp.* **2** 66–96
[10] Choi J W, Hong D, Chang K Y, Chi D P and Lee S 2009 arXiv:0901.1178
[11] Li Q, Li C, Long D Y, Chan W H and Wu C H 2012 *Quantum Inf. Process.* **11** 519–27
[12] Chailloux A, Kerenidis I and Rosgen B 2011 Quantum commitments from complexity assumptions *Automata, Languages and Programming (Lecture Notes in Computer Science* ed L Aceto, M Henzinger and J Sgall vol 6755 (Berlin: Springer) 73–85
[13] Kent A 2012 *Phys. Rev. Lett.* **109** 130501
[14] Aharonov D, Ta-Shma A, Vazirani U and Yao A 2000 Quantum bit escrow *Proc. 32nd Ann. ACM Symp. on Theory of Computing STOC 00* (New York, NY, USA: ACM) 705–14
[15] Ambainis A 2004 *J. Comput. Syst. Sci.* **68** 398–416
[16] Wehner S, Schaffner C and Terhal B M 2008 *Phys. Rev. Lett.* **100** 220502
[17] Schaffner C, Terhal B and Wehner S D C 2009 *Quantum Inf. Comp.* **11–2** 963–96
[18] Wehner S, Curty M, Schaffner C and Lo H K 2010 *Phys. Rev.* A **81** 052336
[19] Konig R, Wehner S and Wullschleger J 2012 *IEEE Trans Inf. Theory* **58** 1962–84
[20] Ng N H Y, Joshi S K, Chen Ming C, Kurtsiefer C and Wehner S 2012 *Nat. Commun.* **3** 1326
[21] Danan A and Vaidman L 2012 *Quantum Inf. Process.* **11** 769–75
[22] Bennett C H and Brassard G 1984 Quantum cryptography: Public key distribution and coin tossing *Proc. IEEE Int. Conf. on Computers Systems and Signal Processing* vol 175 (Bangalore, India) 175–79
[23] Kent A 2011 *New J. Phys.* **13** 113015
[24] Croke S and Kent A 2012 *Phys. Rev.* A **86** 052309
[25] Kaniewski J, Tomamichel M, Hanggi E and Wehner S 2013 *IEEE Trans. Inf. Theory* **59** 4687–99
[26] Lunghi T, Kaniewski J, Bussières F, Houlmann R, Tomamichel M, Kent A, Gisin N, Wehner S and Zbinden H 2013 *Phys. Rev. Lett.* **111** 180504
[27] Liu Y *et al* 2014 *Phys. Rev. Lett.* **112** 010504
[28] Simard J R 2007 Classical and quantum strategies for bit commitment schemes in the two-prover model *MSc Thesis* McGill University
[29] Lunghi T, Kaniewski J, Bussières F, Houlmann R, Tomamichel M, Wehner S and Zbinden H 2015 *Phys. Rev. Lett.* **115** 030502
[30] Ekert A K, Huttner B, Palma G M and Peres A 1994 *Phys. Rev.* A **50** 1047–56
[31] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 *Rev. Mod. Phys.* **74** 145–95
[32] Maurer P C *et al* 2012 *Science* **336** 1283–86
[33] Saeedi K, Simmons S, Salvail J Z, Dluhy P, Riemann H, Abrosimov N V, Becker P, Pohl H J, Morton J J L and Thewalt M L W 2013 *Science* **342** 830–33
[34] Fuchs C A, Gisin N, Griffiths R B, Niu C S and Peres A 1997 *Phys. Rev.* A **56** 1163–72
[35] Williamson M and Vedral V 2003 *J. Mod. Opt* **50** 1989–2011
[36] Dan L, Chang-xing P, Dong-xiao Q, Bao-bin H and Nan Z 2009 A new attack strategy for BB84 protocol based on, Breidbart basis *4th Int. Conf. on Communications and Networking in China, 2009. ChinaCOM 2009* 1–3
[37] Ribordy G, Gisin N, Guinnard O, Stucki D, Wegmuller M and Zbinden H 2004 *J. Mod. Opt* **51** 1381–98
[38] Vedral V 2002 *Rev. Mod. Phys.* **74** 197–234
[39] Goldreich O 2005 *Found. Trends Theor. Comput. Sci.* **1** 1–16

Journal of Physics A: Mathematical and General

On statistical and deterministic quantum teleportation

To cite this article: Vladimir Bozin *et al* 1998 *J. Phys. A: Math. Gen.* **31** 4403

View the article online for updates and enhancements.

You may also like

- To share and not share a singlet: control qubit and nonclassicality in teleportation
Kornikar Sen, Adithi Ajith, Saronath Halder et al.
- Relating EPR steering with the fidelity of quantum teleportation for two- and three-qubit states
Yi Fan, Liang Qiu, Chuanlei Jia et al.
- Versatile and controlled quantum teleportation network
Yao-Yao Zhou, , Peng-Xian Mei et al.

This content was downloaded from IP address 147.91.81.84 on 19/09/2025 at 15:33

J. Phys. A: Math. Gen. **31** (1998) 4403–4435. Printed in the UK PII: S0305-4470(98)88588-0

On statistical and deterministic quantum teleportation

Vladimir Božin†‡, Marina Hruška†, Nikola Paunković† and Fedor Herbut†§

† Faculty of Physics, University of Belgrade, PO Box 368, 11001 Beograd, Yugoslavia
‡ Faculty of Mathematics, University of Belgrade, Studentski TRG 12, Beograd, Yugoslavia
§ Serbian Academy of Sciences and Arts, Knez Mihajlova 35, Beograd, Yugoslavia

Received 20 October 1997

Abstract. We generalize quantum teleportation to, what we call, statistical teleportation utilizing previous results on distant preparation, and on the basic ingredient entities of an entangled composite-system state vector. Our main result is 'the central theorem', establishing a simple necessary and sufficient condition for the crucial entity: the event that the sender of a pure quantum state has to measure in the first step of the two-step (and two-laboratory) teleportation procedure. We derive numerous consequences especially for deterministic teleportation (a special case of statistical teleportation), which is a direct generalization of the known quantum teleportation. Detailed further generalization to proper and improper mixtures is investigated. Finally, it is shown that extension to teleportation with nonlinear distant preparation is not possible unless the idea of teleportation is essentially changed.

1. Introduction

Entangled composite-system state vectors are known to display nonclassical statistical correlations between the subsystems (so-called nonseparability (D'Espagnat 1976)). A major breakthrough in the physical understanding of nonseparability was made by Schrödinger (1936) when he discovered distant preparation and when he obtained the first results on its amazing scope.

Entanglement theory, also known as distant correlations theory, was further developed in great detail and generality along the lines of Schrödinger's approach by one of us (FH) and Vujičić (1976) (Vujičić and Herbut 1984, 1988). It was done in the framework of antilinear Hilbert–Schmidt operators (which have not found sufficient application in the further literature on distant correlations) and distant measurement (which is not particularly relevant for teleportation). For the readers' convenience, most of the previous results on distant correlations that we find relevant for teleportation are restated in this article in a new and simplified way. Since they are rather elementary and not surprising, the proofs are omitted.

In understanding entanglement, the recent discovery of teleportation by Bennett *et al* (1993) seems to be one of the major events. It is receiving continued attention (e.g. Popescu 1994, Bennett 1995, Bennett *et al* 1996, Nielsen and Caves 1997, Moussa 1997 etc).

Investigating how wide the choice of the entities used in quantum teleportation is, i.e. what features these entities should have in order to make quantum teleportation possible, we find it appropriate to introduce the concept of *statistical teleportation* (see section 3). It corresponds to a particular event that is measured as an observable in the first step. Its occurrence (result 1) takes place with a probability, or, equivalently, in a subensemble. It is also the basic building block of deterministic teleportation (see (9) below).

0305-4470/98/194403+33$19.50 © 1998 IOP Publishing Ltd

Statistical teleportation is a two-step and two-laboratory phenomenon. The first step takes place in the nearby laboratory in the way of a measurement that via *distant preparation* of a suitable state vector, and a classical message, brings about a change in the distant laboratory. The second step consists of the application of a suitable unitary operation in the distant laboratory converting the distantly prepared state vector into the teleported one.

Investigation of the first step of statistical teleportation is not only an application, but also a further development of distant preparation theory: it is elaboration of the special case when distant preparation is a linear map.

Some attempts at practical application of deterministic or perfect teleportation actually fall back on statistical teleportation and achieve only a 25%, or at best a 50%, absolute efficiency (Braunstein and Mann 1995). Hence, in applications, statistical teleportation can be expected to play an important role both as a constituent of perfect, or of partially realized or imperfect, teleportation or as an outright competitor to perfect teleportation.

In order to gain a balanced view about the potential importance of the concept of statistical teleportation that is being proposed, one should also be aware of the fact that recently, in the context of the notion of interaction-free measurement, the idea of statistical measurement appeared in the literature (Elitzur and Vaidman 1993, Kwiat *et al* 1995, 1996). Namely, in individual cases the measurement at issue may or may not succeed; it takes place with a probability. Ensemblewise speaking, we have measurement in a subensemble, just like the case with teleportation when it is statistical.

A summary of distant-preparation theory is presented in section 2. In section 3 the concept of statistical teleportation is defined and its relation to deterministic teleportation is clarified. A summary on the basic ingredients of a composite-system state vector is given in section 4. The central result on statistical teleportation is derived in section 5. Its immediate consequences are derived in section 6. Section 7 is devoted to a detailed discussion of deterministic teleportation. Generalization to proper and improper mixtures is performed in sections 8 and 9 respectively. Finally, concluding remarks are given in section 10.

Since we shall be dealing with entities containing redundancies, we shall extract the relevant parts (relevant entities) from them. Whenever possible we utilize a simpler notation (and simpler term) for the relevant entity. In particular, the redundant entities will be denoted with bars, and the relevant ones without them.

2. Partial scalar product and distant preparation

In this section we restate some relevant basic results on distant preparation developed in the previously mentioned articles.

Distant preparation requires an entangled state vector of a composite system consisting of two subsystems. It comes about when one chooses one of the subsystems to perform direct subsystem measurement on it. We call the directly measured subsystem the nearby one, and we denote it by 'n'. In distant preparation one studies the consequences of the direct subsystem measurement on the state of the other subsystem. We call the latter the distant subsystem and we denote it by 'd'.

To obtain the mentioned consequences, it is important that, during the direct measurement on the nearby subsystem, neither the measuring apparatus nor this subsystem must interact with the distant subsystem. The latter changes its quantum-mechanical state due to the very procedure described, and this is called *distant preparation*. Most often the subsystems are far apart or distant from each other so that the direct subsystem measurement on the nearby subsystem can be easily performed without influencing dynamically the state

of the distant subsystem.

We denote the state spaces (Hilbert spaces) of the nearby and the distant subsystems by $\mathcal{H}_n$ and $\mathcal{H}_d$ respectively.

Lemma 1. Let $|\chi\rangle_{nd}$ $(\in (\mathcal{H}_n \otimes \mathcal{H}_d))$ and $|\phi\rangle_n$ $(\in \mathcal{H}_n)$ be two nonzero vectors. Let $|\chi\rangle_{nd} = \sum_{i=1}^{I} \alpha_i |i\rangle_n |i\rangle_d$ ('I' finite or $\infty, \forall i : \alpha_i \in \boldsymbol{C}$) be an expansion of $|\chi\rangle_{nd}$ in terms of uncorrelated vectors. Then the vector

$$\sum_{i=1}^{I} (\alpha_i \langle\phi|_n |i\rangle_n) |i\rangle_d$$

whereby $(\langle\phi|_n |i\rangle_n)$ we denote the number obtained by ordinary scalar multiplication in $\mathcal{H}_n$, is an element of $\mathcal{H}_d$, and it is one and the same whatever the mentioned expansion of $|\chi\rangle_{nd}$ in uncorrelated vectors.

Definition 1. The so-called *partial scalar product*, denoted by $\langle\phi|_n|\chi\rangle_{nd}$ and belonging to $\mathcal{H}_d$, is evaluated by expanding $|\chi\rangle_{nd}$ in terms of uncorrelated vectors in an arbitrary way (cf lemma 1), and by using the ordinary scalar product in $\mathcal{H}_n$. If we have, for example $|\chi\rangle_{nd} = \sum_{i=1}^{I} \alpha_i |i\rangle_n |i\rangle_d$, then, by definition,

$$\langle\phi|_n |\chi\rangle_{nd} \equiv \sum_{i=1}^{I} (\alpha_i \langle\phi|_n |i\rangle_n) |i\rangle_d. \tag{1}$$

Corollary 1. One has

$$\langle\lambda|_d (\langle\phi|_n |\chi\rangle_{nd}) = (\langle\phi|_n \langle\lambda|_d) |\chi\rangle_{nd}.$$

(Note that the LHS is a scalar product in $\mathcal{H}_d$, and the RHS is one in $(\mathcal{H}_n \otimes \mathcal{H}_d)$.)

Corollary 2. Partial scalar multiplication is continuous and linear in the composite-system vector.

Corollary 3. If A_n is a bounded Hermitian operator in $\mathcal{H}_n$, then we have the following equality of partial scalar products

$$\langle\phi|_n ((A_n \otimes I_d) |\chi\rangle_{nd}) = (\langle\phi|_n A_n) |\chi\rangle_{nd}$$

where I_d is the identity operator in $\mathcal{H}_d$.

Corollary 4. If A_d is a bounded linear operator in $\mathcal{H}_d$, then

$$A_d (\langle\phi|_n |\chi\rangle_{nd}) = \langle\phi|_n ((I_n \otimes A_d) |\chi\rangle_{nd})$$

always, where I_n is the identity operator in $\mathcal{H}_n$.

Corollary 5. Let $|\phi\rangle_n \in \mathcal{H}_n$, $\langle\phi|_n |\phi\rangle_n = 1$. Then

$$(|\phi\rangle_n \langle\phi|_n \otimes I_d) |\chi\rangle_{nd} = |\phi\rangle_n \otimes (\langle\phi|_n |\chi\rangle_{nd}). \tag{2}$$

In this outline we deal with *distant preparation* under two restrictions. We assume that the composite system is in a pure state, described by a state vector $|\chi\rangle_{nd}$ for example and that it is an elementary event (or atom), $|\phi\rangle_n \langle\phi|_n$ for example, that occurs or not on the nearby subsystem.

If the observable (projector) $|\phi\rangle_n \langle\phi|_n$ (or $(|\phi\rangle_n \langle\phi|_n \otimes I_d)$) is measured and if one obtains 1 as the result of the measurement, then, one speaks of the *occurrence of the event* at issue (see Gudder 1979). The result 0 is referred to as nonoccurrence.

Switching over from the more usual individual-system language to the (equivalent) ensemble one, one can say that after the measurement of $(|\phi\rangle_n \langle\phi|_n \otimes I_d)$ in $|\chi\rangle_{nd}$ we leave out all those distant subsystems on the nearby partners of which $|\phi\rangle_n \langle\phi|_n$ has not occurred.

The remaining subensemble of distant subsystems, on the nearby partners of which $|\phi\rangle_n\langle\phi|_n$ has occurred, is, by definition, *the distantly prepared subensemble* (or state).

The fraction of distant subsystems that goes into the distantly prepared subensemble is obviously equal to the probability of the occurrence of $(|\phi\rangle_n\langle\phi|_n \otimes I_d)$ in $|\chi\rangle_{nd}$.

The distantly prepared state can be easily evaluated if one confines oneself to ideal measurement. However, before we state this result, we introduce the concept and evaluation of the state operator of a subsystem.

Definition 2. If a composite system is in the state given by the state vector $|\chi\rangle_{nd}$, then the state of the distant subsystem is given by the reduced statistical operator (state operator)

$$\rho_d \equiv \mathrm{Tr}_n\, |\chi\rangle_{nd}\langle\chi|_{nd} \equiv \sum_i (\langle\phi_i|_n|\chi\rangle_{nd})(\langle\chi|_{nd}|\phi_i\rangle_n) \tag{3}$$

where Tr_n denotes the partial trace in $\mathcal{H}_n$, and $\{|\phi_i\rangle_n : \forall i\}$ is a complete orthonormal (ON) basis in $\mathcal{H}_n$. The definition and evaluation of the nearby-subsystem state ρ_n is symmetrical to (3) with respect to interchange of n and d.

Taking the matrix representation of ρ_d in a complete ON basis in $\mathcal{H}_d$, the RHS of (3) is, due to corollary 1, seen to reduce to the usual definition of the partial trace, which is known to be independent both of the choice of the basis $\{|\phi_i\rangle_n : \forall i\}$ in $\mathcal{H}_n$ (cf (3)) and the mentioned basis in $\mathcal{H}_d$.

Corollary 6. For any bounded Hermitian operators A_d and B_d one has

$$A_d[\mathrm{Tr}_n(|\chi\rangle_{nd}\langle\chi|_{nd})]B_d = \mathrm{Tr}_n[(I_n \otimes A_d)(|\chi\rangle_{nd}\langle\chi|_{nd})(I_n \otimes B_d)].$$

And symmetrically, for any bounded Hermitian operators A_n and B_n:

$$A_n[\mathrm{Tr}_d(|\chi\rangle_{nd}\langle\chi|_{nd})]B_n = \mathrm{Tr}_d[(A_n \otimes I_d)(|\chi\rangle_{nd}\langle\chi|_{nd})(B_n \otimes I_d)].$$

Proposition 1 (distant preparation by ideal subsystem measurement). Let $|\phi\rangle_n$ $(\in \mathcal{H}_n)$ and $|\chi\rangle_{nd}$ $(\in (\mathcal{H}_n \otimes \mathcal{H}_d))$ be state vectors, and let us write

$$\langle\phi|_n|\chi\rangle_{nd} = w^{1/2}|\omega\rangle_d \tag{4}$$

where w is by definition the square of the norm of the vector on the LHS. If the event $(|\phi\rangle_n\langle\phi|_n \otimes I_d)$ occurs in *ideal* nearby-subsystem measurement in the state $|\chi\rangle_{nd}$ of the composite system, then the probability of the occurrence is w, and the distantly prepared state vector is $|\omega\rangle_d$.

Proof. The required probability p is, as well known,

$$p = \langle\chi|_{nd}(|\phi\rangle_n\langle\phi|_n \otimes I_d)|\chi\rangle_{nd}$$

which, utilizing idempotency and (2) twice, gives

$$p = \|\langle\phi|_n|\chi\rangle_{nd}\|^2 = w$$

as claimed.

Since in ideal measurement the change of state due to occurrence is described by the Lüders formula (Messiah 1961, Lüders 1951), the occurrence of $(|\phi\rangle_n\langle\phi|_n \otimes I_d)$ in $|\chi\rangle_{nd}$ changes this composite-system state vector into the state vector

$$|\bar{\chi}\rangle_{nd} \equiv w^{-1/2}(|\phi\rangle_n\langle\phi|_n \otimes I_d)|\chi\rangle_{nd}.$$

The distant subsystem is, then, described by the reduced statistical operator (cf (3))

$$\bar{\rho}_d \equiv \mathrm{Tr}_n\, |\bar{\chi}\rangle_{nd}\langle\bar{\chi}|_{nd}$$

which, after substitution of the preceding relation, using (2), and finally of (4), becomes

$$\bar{\rho}_d = \mathrm{Tr}_n(|\phi\rangle_n\langle\phi|_n \otimes |\omega\rangle_d\langle\omega|_d) = |\omega\rangle_d\langle\omega|_d$$

as claimed. □

As is well known, three requirements are relevant for individual-system measurements. (i) In each measurement a result (characteristic value) is obtained, and their relative frequencies, in the limit of infinitely many measurements, reproduce the probability values predicted by quantum mechanics. (ii) The results are repeatable, i.e. an immediately repeated measurement of the same observable on the same individual system gives the same result with certainty. (iii) Possible results on compatible observables are preserved in each individual-system measurement.

If all three requirements are satisfied, we have ideal measurement (cf Herbut 1974, theorem 1). If the third is not, but the first two are, then we are dealing with nonideal first-kind or repeatable measurements (Pauli 1933, Busch *et al* 1991). Finally, if only requirement (i) is valid, one has second-kind or nonrepeatable measurement (ibid).

If one has in mind nonideal first-kind or second-kind measurement, then the probability is the same as in ideal measurement because it does not depend on the kind of measurement performed. But, one wonders if a similar statement is true about the distantly prepared subensemble: is it still describable by a state operator? If yes, can this be evaluated? An answer to the latter question may be burdened by the fact that there is no change-of-state formula in nonideal measurement.

Our chance lies in the fact that the concept of 'state' serves to provide us with probabilities, and it is determined by the totality of the latter. Hence, we have to convert our change-of-state problem into the language of mere probabilities.

In an attempt to generalize distant preparation from ideal to *any kind of individual-system measurement*, it seems reasonable to base our further considerations on the following *two physical assumptions*.

(i) The probability of occurrence of two compatible events in immediate succession (i.e. when the time interval between the occurrences tends to zero) equals the probability of coincidence, i.e. of joint measurement of the two compatible events. And this is valid for any kind of measurement.

(ii) If $|\lambda\rangle_d$ is an arbitrary state vector in $\mathcal{H}_d$, there exists a probability $p(\lambda)$ for the occurrence of $|\lambda\rangle_d\langle\lambda|_d$ in the distantly prepared state by any measurement. (This probability is conditional by the very definition of the distantly prepared state.)

Now, we can prove the basic relevant result of this section.

Theorem 1. Let $|\phi\rangle_n$ and $|\chi\rangle_{nd}$ be state vectors. We assume that in an arbitrary measurement of the event $(|\phi\rangle_n\langle\phi|_n \otimes I_d)$ in the state $|\chi\rangle_{nd}$ the occurrence has a positive probability w. The occurrence brings *ipso facto* the distant subsystem into the state $|\omega\rangle_d$ defined by (4). Besides, also the probability w is given by (4) (just like in ideal measurement).

Proof. We apply the above two assumptions to the immediate succession of the events $(|\phi\rangle_n\langle\phi|_n \otimes I_d)$ and $(I_n \otimes |\lambda\rangle_d\langle\lambda|_d)$ in the state $|\chi\rangle_{nd}$:

$$wp(\lambda) = \langle\chi|_{nd}(|\phi\rangle_n\langle\phi|_n \otimes |\lambda\rangle_d\langle\lambda|_d)|\chi_{nd}\rangle. \tag{5}$$

(Note that on the LHS we have the usual conditional-probability factorization.)

The probability $p(\lambda)$ is determined by the rest of the entities in (5), and all these, being probabilities, are independent of the kind of measurement. Hence, the probability $p(\lambda)$ is in any measurement the same as in ideal measurement. In the latter we know (see

proposition 1) that $p(\lambda) = |\langle\omega|_d|\lambda\rangle_d|^2$, i.e. that the distantly prepared state is described by $\rho_d = |\omega\rangle_d\langle\omega|_d$ (cf (4)). Hence, this is true for all measurements. □

We have thus been able to generalize the physical meaning of $|\omega\rangle_d$ defined in (4) from ideal to any kind of measurement on account of the above physical assumptions. Bell and Nauenberg (1966) have made assumption (i), but only for ideal measurement (on a noncomposite system), and they have derived the Lüders formula (Messiah 1961, Lüders 1951). We have extended their assumption to all kinds of measurement. This seems justified because probabilities in quantum mechanics do not depend on the kind of measurement performed.

3. Definition of statistical teleportation

The investigation that follows will be based on relation (4) evaluating the distantly prepared state vector. Let us introduce the concepts required for a definition of statistical teleportation.

The system on which quantum teleportation takes place consists of three subsystems. We denote them by 1, 2, and 3. There are two laboratories. The experimenter in the nearby one, who is the sender of a quantum-state vector, has the nearby subsystem $(1 + 2)$ at his disposal, whereas the experimenter in the distant laboratory, who is the receiver, is able to act on the state of the distant subsystem 3.

The most important entity for teleportation is a given entangled state vector $|\bar{\Psi}\rangle_{23}$ of the composite system $(2 + 3)$. It plays the role of a sort of a bridge or a specific quantum information channel because, as stated, subsystem 2 is in the nearby laboratory, and subsystem 3 is in the distant one. We call $|\bar{\Psi}\rangle_{23}$ *the bridge state vector*.

In the nearby laboratory one measures a suitable observable, an elementary event (synonym: an atom) $|\bar{a}\rangle_{12}\langle\bar{a}|_{12}$, which is determined by a state vector $|\bar{a}\rangle_{12}$ that we call the *atomic state vector*. This measurement is actually performed in the three-subsystem composite-system state described by $|\psi\rangle_1|\bar{\Psi}\rangle_{23}$, where $|\psi\rangle_1$ is a state vector for subsystem 1 that is to be teleported. As a consequence of the occurrence of the event $(|\bar{a}\rangle_{12}\langle\bar{a}|_{12} \otimes I_3)$ in its measurement (I_3 being the identity operator in $\mathcal{H}_3$), that occurs with some positive probability $\bar{w}$, a distantly prepared state $|\psi\rangle'_3$ appears in the distant laboratory. It is evaluated by taking the partial scalar product

$$\langle\bar{a}|_{12}|\psi\rangle_1|\bar{\Psi}\rangle_{23} = (\bar{w})^{1/2}|\psi\rangle'_3 \tag{6a}$$

(see proposition 1 and theorem 1).

The experimenter in the distant laboratory has an operation $\bar{\mathcal{U}}_3$ at his disposal, which is suitably chosen for (or associated with) the atomic state vector $|\bar{a}\rangle_{12}$. If the crucial event $(|\bar{a}\rangle_{12}\langle\bar{a}|_{12} \otimes I_3)$ does occur, a classical message (a phone call, for example) goes from the nearby laboratory to the distant one. Then in the latter the mentioned operation $\bar{\mathcal{U}}_3$ is applied to the distantly prepared state $|\psi\rangle'_3$, and it is thus converted into the teleported state $|\psi\rangle_3$.

More precisely, let $\boldsymbol{S}_1$ be some linear manifold of first-subsystem states. Then, for each $|\psi\rangle_1 \in \boldsymbol{S}_1$,

$$|\psi\rangle_3 = \bar{\mathcal{I}}_{31}|\psi\rangle_1 \tag{7}$$

where '$\bar{\mathcal{I}}_{31}$' is the well known unitary isomorphism mapping $\mathcal{H}_1$ onto $\mathcal{H}_3$ and giving *physically the same* state, or it is some other fixed isomorphism of $\mathcal{H}_1$ onto $\mathcal{H}_3$. (If the latter is the case, then we have some kind of generalized teleportation, but the theory is formally the same). If $\bar{\mathcal{I}}_{31}$ is the 'same-state' isomorphism, then subsystems 1 and 3 have

to be of the same physical nature, of course. If $\bar{\mathcal{I}}_{31}$ is more general, $\mathcal{H}_1$ and $\mathcal{H}_3$ have to be only isomorphic. We denote by '$\mathcal{I}_{31}$' the map obtained by restricting the domain of $\bar{\mathcal{I}}_{31}$ to $\boldsymbol{S}_1$.

One can make the following summary.

Definition 3. We call a two-step procedure in two distantly separated laboratories that is performed with the purpose to transmit an arbitrary state vector $|\psi\rangle_1$ from some linear manifold $\boldsymbol{S}_1$ in $\mathcal{H}_1$ into the physically identical (or isomorphic) state vector $|\psi\rangle_3$ in $\mathcal{H}_3$ *statistical teleportation* if the following requirements are satisfied:

(i) The first step is distant preparation, we denote it by $\mathcal{U}_{31}$, taking state vectors from $\boldsymbol{S}_1$ into state vectors in $\mathcal{H}_3$. It is performed in the nearby laboratory by measuring a suitable observable, an event of the form $(|\bar{a}\rangle_{12}\langle\bar{a}|_{12} \otimes I_3)$, $|\bar{a}\rangle_{12} \in (\mathcal{H}_1 \otimes \mathcal{H}_2)$, in the state $|\Psi\rangle_1|\bar{\Psi}\rangle_{23}$, where $|\bar{\Psi}\rangle_{23}$ is a given entangled state vector in $(\mathcal{H}_2 \otimes \mathcal{H}_3)$, shared between the two laboratories (subsystem 2 being in the nearby laboratory and subsystem 3 in the distant laboratory). The suitable chosen state vector $|\bar{a}\rangle_{12}$ is called atomic.

(ii) Information about occurrence or nonoccurrence of the mentioned event is transmitted by a classical channel from the nearby to the distant laboratory. In case of occurrence, the second step of teleportation, a unitary operation in $\mathcal{H}_3$, transforms the distantly prepared state of subsystem 3 into the physically identical (or isomorphic) teleported state vector.

We denote the linear manifold spanned by the range of $\mathcal{U}_{31}$, i.e. $\mathcal{U}_{31}(\boldsymbol{S}_1)$ by $\boldsymbol{S}'_3$, the restriction of operation $\bar{\mathcal{U}}_3$ to $\boldsymbol{S}_3$ by $\mathcal{U}_3$, and, finally, the linear manifold $\mathcal{U}_3(\boldsymbol{S}'_3)$ by $\boldsymbol{S}_3$.

The two-step and two-laboratory procedure at issue is described by the following chain of operations

$$\mathcal{I}_{31} = \mathcal{U}_3 \circ \mathcal{U}_{31} \tag{8}$$

where '$\circ$' denotes 'after' reading from left to right, as displayed on the commuting figure 1. (So far (8) is confined to state vectors in $\boldsymbol{S}_1$.)

The term 'statistical teleportation' is suggested having an individual (1+2+3) composite system in mind. One may also use the synonym 'teleportation with a probability'. If one envisages an ensemble of $(1 + 2 + 3)$ systems of the kind described, one may use the synonym 'teleportation in a subensemble'.

We now *require* $\mathcal{U}_3$ to be *linear*, which, in view of the fact that it preserves the norm, is equivalent to requiring $\mathcal{U}_3$ to be a unitary isomorphism mapping $\boldsymbol{S}'_3$ onto $\boldsymbol{S}_3$. On account of relation (8), also $\mathcal{U}_{31}$ is then linear and hence a unitary isomorphism taking $\boldsymbol{S}_1$ onto $\boldsymbol{S}'_3$. (This is where elaboration of distant preparation $\mathcal{U}_{31}$ comes into play.)

The above requirement of linearity of $\mathcal{U}_3$ (and of $\mathcal{U}_{31}$) is crucial for our theory. At the end of this article (cf section 10.1) this requirement is investigated, and it is proved that it is a necessary consequence of some very basic assumptions.

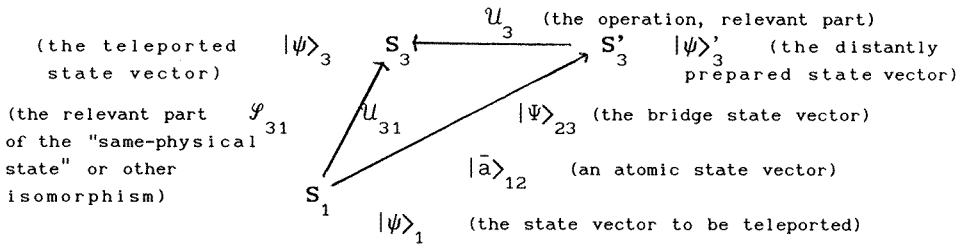


Figure 1.

Evidently, the distant experimenter must interact very nontrivially with subsystem 3 to bring about the action of $\mathcal{U}_3$. This interaction, whatever its form, is a part of the spontaneous evolution of the distant laboratory. Unitary evolution is confined, as well known, to dynamically isolated systems, and subsystem 3 is not at all isolated.

A way out of this apparent paradox lies in the fact that the instruments with which the distant experimenter acts on subsystem 3 are macroscopic bodies. Thus, they are tremendously large compared to subsystem 3, which is thought of as a microscopic system. Hence, the feedback (or reaction) on the instruments can be neglected in a very good approximation. This makes subsystem 3 'move in an external potential'.

We have seen that statistical teleportation is an intricate process that has distant preparation at its heart. *Deterministic* (or perfect) *teleportation*, invented by Bennett *et al* (1993), is slightly more complex than statistical teleportation, and the latter is the essential ingredient of the former.

Namely, one has a complete observable in $\mathcal{H}_{12}$ (i.e. one with all characteristic values nondegenerate) all characteristic vectors of which are atomic state vectors:

$$A_{12} \equiv \sum_{m=1}^{M} a_m |\bar{a}^{(m)}\rangle_{12}\langle\bar{a}^{(m)}|_{12} \qquad (m \neq m' \Rightarrow a_m \neq a_{m'}) \tag{9}$$

where $M \equiv \dim(\mathcal{H}_1) \times \dim(\mathcal{H}_2)$.

The distant experimenter has an operation $\bar{\mathcal{U}}_3^{(m)}$ associated with each of the atomic state vectors $|\bar{a}^{(m)}\rangle_{12}$ at his disposal. After the measurement of $(A_{12} \otimes I_3)$ in the state $|\psi\rangle_1|\bar{\Psi}\rangle_{23}$ a 'phone call' informs the distant experimenter which result a_m has occurred. He applies the corresponding operation $\bar{\mathcal{U}}_3^{(m)}$, and thus he necessarily obtains one and the same teleported state $|\psi\rangle_3$ $(\equiv \mathcal{I}_{31}|\psi\rangle_1)$ whatever the classical message. This is why we speak of 'deterministic' teleportation, or teleportation with certainty, in this case. (The ensemble point of view is unnecessary here.)

It is our first aim to find all atomic state vectors $|\bar{a}\rangle_{12}$, i.e. state vectors which make statistical teleportation possible when a bridge state vector $|\bar{\Psi}\rangle_{23}$ and a linear manifold $\boldsymbol{S}_1'(\subseteq \mathcal{H}_1)$ are given. The operation $\mathcal{U}_3$ is, then, determined by $\mathcal{U}_{31}$ (cf (8)), which is, in turn, determined by $|\bar{a}\rangle_{12}$ via (6*a*), rewritten as

$$\mathcal{U}_{31}|\psi\rangle_1 = |\psi\rangle_3'. \tag{6b}$$

But before we proceed with our theory of statistical teleportation, we expound further the 'tools' of distant preparation theory because we need a precise statement on the scope of distant preparation (to see how wide the linear manifold $\boldsymbol{S}_3$ can turn out to be).

4. The basic ingredient entities of a composite-system state vector

In this section we restate the relevant part of a distant correlations theory from previous work (Herbut and Vujičić 1976, Vujičić and Herbut 1984 and 1988). We resume the notation of section 2 calling the two subsystems the 'nearby' and the 'distant' one.

The subsystem state operators and the correlation operator, the basic ingredient entities of a composite-system state vector $|\chi\rangle_{nd}$ $(\in (\mathcal{H}_n \otimes \mathcal{H}_d))$, are the tools in terms of which our theory of teleportation is going to be developed.

Lemma 2. Let $|\chi\rangle_{nd}$ be a composite-system state vector, and $\{|\phi_k\rangle_n : \forall k\}$ a complete ON basis in $\mathcal{H}_n$. Then the following claims are valid.

(i) There exists a set of vectors $\{|\omega_k'\rangle_d : \forall k\}$ $(\subseteq \mathcal{H}_d)$ such that

$$|\chi\rangle_{nd} = \sum_k |\phi_k\rangle_n|\omega_k'\rangle_d. \tag{10a}$$

(ii) The vectors $|\omega'_k\rangle_d$, called generalized expansion coefficients, are unique and each depends only on the corresponding subsystem basis vector $|\phi_k\rangle_n$ and on $|\chi\rangle_{nd}$.

(iii) The generalized expansion coefficients can be evaluated as partial scalar products:

$$\forall k: \qquad |\omega'_k\rangle_d = \langle\phi_k|_n|\chi\rangle_{nd}. \tag{10b}$$

The subsystem state operators (reduced statistical operators) ρ_n and ρ_d of a given composite-system state vector $|\chi\rangle_{nd}$ were defined in definition 2.

Lemma 3. A generalized expansion coefficient $|\omega'_k\rangle_d$ in (10*a*) is zero if and only if the corresponding vector $|\phi_k\rangle_n$ belongs to the null space of ρ_n. Equivalently, one has an expansion in a subsystem ON subbasis $\{|\phi_k\rangle_n : \forall k\}$ of the form (10*a*) if and only if the subspace spanned by this subbasis contains the range of ρ_n.

Corollary 7. If a subsystem ON subbasis $\{|\phi_k\rangle_n : \forall k\}$ such that $|\chi\rangle_{nd}$ can be expanded in it (cf lemma 3) is given, and in the corresponding generalized expansion coefficients $|\omega'_k\rangle_d$ in (10*a*) the norms '$w_k^{1/2}$' are explicitly displayed

$$\forall k: |\omega'_k\rangle_d \equiv w_k^{1/2}|\omega_k\rangle_d \tag{11a}$$

then one has the following (mathematical) decomposition of the state operator ρ_d of the distant subsystem into pure states:

$$\rho_d = \sum_k w_k|\omega_k\rangle_d\langle\omega_k|_d. \tag{11b}$$

From the physical point of view, decomposition (11*b*) is a potential, not an actual, one because, taking the ensemble point of view, ρ_d cannot be actually thought of as decomposed into distinct subensembles (corresponding to the terms in (11*b*)) in view of the homogeneity of the ensemble $|\chi\rangle_{nd}\langle\chi|_{nd}$ of composite systems. But the (nonselective) measurement of a complete nearby-subsystem observable

$$(B_n \otimes I_d) = \sum_k b_k(|\phi_k\rangle_n\langle\phi_k|_n \otimes I_d) \qquad k \neq k' \Rightarrow b_k \neq b_{k'}$$

(the $|\phi_k\rangle_n$ same as in lemma 3) in the state $|\chi\rangle_{nd}$ brings about (11*b*) as an actual decomposition. This is an immediate consequence of theorem 1. It is called *distant state operator decomposition* (or empirically: *distant ensemble decomposition*). It is the nonselective counterpart of distant preparation. (The latter is a notion connected with the selective measurement of one characteristic value b_k of $(B_n \otimes I_d)$.)

Lemma 4. Expansion (10*a*) in a subsystem ON subbasis is biorthogonal, i.e. also the set $\{|\omega'_k\rangle_d : \forall k\}$ of the generalized expansion coefficients is a set of orthogonal vectors, if and only if $\{|\phi_k\rangle_n : \forall k\}$ is a characteristic subbasis of ρ_n.

Definition 4. If one writes a biorthogonal expansion of $|\chi\rangle_{nd}$ in terms of ON subbases $\{|\phi_i\rangle_n : \forall i\}$ and $\{|\omega_i\rangle_d : \forall i\}$ for the two subsystems

$$|\chi\rangle_{nd} = \sum_i r_i^{1/2}|\phi_i\rangle_n|\omega_i\rangle_d \tag{12}$$

with positive $r_i^{1/2}$, then one says that a *Schmidt biorthogonal expansion* of $|\chi\rangle_{nd}$ is given. The vectors $(r_i^{1/2}|\omega_i\rangle_d)$ are the generalized expansion coefficients in (10*a*), and $r_i^{1/2}$ are their norms (cf (11*a*)).

It is obvious from lemma 4 and definition 4 that every composite-system state vector $|\chi\rangle_{nd}$ can be written as a Schmidt biorthogonal expansion, and that the latter is, in general, not unique as far as the basis vectors are concerned.

Corollary 8. If a Schmidt biorthogonal expansion (12) is given, one can read off spectral forms of ρ_n and ρ_d. They are given by the relations:

$$\rho_n = \sum_i r_i |\phi_i\rangle_n \langle\phi_i|_n \qquad \forall i : r_i > 0 \tag{13a}$$

$$\rho_d = \sum_i r_i |\omega_i\rangle_d \langle\omega_i|_d \qquad \forall i : r_i > 0. \tag{13b}$$

(Notice that the positive characteristic values of ρ_n and ρ_d necessarily coincide.)

Remark 1. In view of (13*a*) ((13*b*)), one can see that the linear manifold spanned by the ON subbasis $\{|\phi_i\rangle_n : \forall i\}$ (by $\{|\omega_i\rangle_d : \forall i\}$) in the Schmidt biorthogonal expansion (12) is the *range* $R(\rho_n)$ ($R(\rho_d)$). The subspace spanned by the same subbasis is the (topologically) closed range $\bar{R}(\rho_n)$ ($\bar{R}(\rho_d)$). If the subbasis is infinite, the range is a proper subset of the closed range. Otherwise, they coincide.

Definition 5. Let $|\chi\rangle_{nd}$ be given as a Schmidt biorthogonal expansion (12). Let us denote by $\mathcal{U}_a$ and call the correlation operator (implied by $|\chi\rangle_{nd}$) the antiunitary (antilinear unitary) isomorphism mapping $\bar{R}(\rho_n)$ onto $\bar{R}(\rho_d)$ so that

$$\forall i : \mathcal{U}_a |\phi_i\rangle_n = |\omega_i\rangle_d. \tag{14a}$$

Remark 2. Definition (14*a*) is evidently equivalent to

$$\mathcal{U}_a = \sum_i |\omega_i\rangle_d K \langle\phi_i|_n \tag{14b}$$

where 'K' is the operation of complex conjugation. Thus,

$$\forall |\lambda\rangle_n \ (\in \bar{R}(\rho_n)): \ \mathcal{U}_a|\lambda\rangle_n = \sum_i (\langle\phi_i|_n|\lambda\rangle_n)^* |\omega_i\rangle_d.$$

Lemma 5. Though a Schmidt biorthogonal expansion (12) of a given composite-system state vector $|\chi\rangle_{nd}$ is not unique (the characteristic subbases displayed in it are not unique, though the spectra of the subsystem state operators always are), the correlation operator $\mathcal{U}_a$ is always *uniquely* implied by $|\chi\rangle_{nd}$.

Remark 3. Evidently, one can rewrite (12) in any of the following two forms:

$$|\chi\rangle_{nd} = \sum_i r_i^{1/2} |\phi_i\rangle_n (\mathcal{U}_a |\phi_i\rangle_n)_d \tag{15a}$$

$$|\chi\rangle_{nd} = \sum_i r_i^{1/2} (\mathcal{U}_a^{-1} |\omega_i\rangle_d)_n |\omega_i\rangle_d. \tag{15b}$$

Corollary 9. A given $|\chi\rangle_{nd}$ implies its ρ_n, ρ_d, and $\mathcal{U}_a$. But also vice versa, ρ_n, and $\mathcal{U}_a$, or, alternatively, ρ_d and $\mathcal{U}_a^{-1}$, determine $|\chi\rangle_{nd}$.

Corollary 10. If Q_n and Q_d are the *range projectors* of the respective subsystem state operators ρ_n and ρ_d of a given $|\chi\rangle_{nd}$, then always

$$|\chi\rangle_{nd} = (Q_n \otimes I_d)|\chi\rangle_{nd} = (I_n \otimes Q_d)|\chi\rangle_{nd}.$$

Definition 6. We call the number N of terms in the Schmidt canonical form (12) of a given $|\chi\rangle_{nd}$, or equivalently, the number of dimensions of the (necessarily equally dimensional) ranges $R(\rho_p)$, $p = n, d$, *the degree of entanglement* in $|\chi\rangle_{nd}$. If N is finite, we say that $|\chi\rangle_{nd}$ is finitely entangled; otherwise, we speak of infinite entanglement.

Definition 7. If a given composite-system state vector $|\chi\rangle_{nd}$ is finitely entangled with the degree N, and if $\forall i : r_i = 1/N$ (cf (12)), then one says that the vector is *maximally entangled*.

Definition 8. Let $\boldsymbol{S}_n$ be a subspace of $\mathcal{H}_n$ and let Q_n be the range projector of ρ_n of a given composite-system state vector $|\chi\rangle_{nd}$ (cf (12) and (13*a*)). If $R(Q_n) \subseteq \boldsymbol{S}_n$, then we say that $|\chi\rangle_{nd}$ is *within* $\boldsymbol{S}_n$. If, in particular, $R(Q_n) = \boldsymbol{S}_n$, then $|\chi\rangle_{nd}$ is said to be *complete in* $\boldsymbol{S}_n$. (If $R(Q_n) \subset \boldsymbol{S}_n$, then the vector is incomplete in $\boldsymbol{S}_n$). These definitions are valid symmetrically for the distant subsystem.

Remark 4. Every state vector $|\chi\rangle_{nd}$ is complete in the closed ranges $\bar{R}(\rho_n)$ $(= R(Q_n))$ and $\bar{R}(\rho_d)$ $(= R(Q_d))$, where Q_n and Q_d are the respective range projectors of ρ_n and ρ_d.

In the main result of this section, the square root of the subsystem state operator $\rho_d^{1/2}$ is going to play a decisive role. It is desirable to gain some knowledge on it.

Remark 5. In case of infinite entanglement, the spectral form (13*b*) implies, as easily seen:

$$R(\rho_d) = \{|\lambda\rangle_d : |\lambda\rangle_d \in \mathcal{H}_d, \sum_i (|\omega_i|_d|\lambda\rangle_d|^2/r_i^2) < \infty\} \tag{16}$$

$$R(\rho_d^{1/2}) = \{|\lambda\rangle_d : |\lambda\rangle_d \in \mathcal{H}_d, \sum_i (|\langle\omega_i|_d|\lambda\rangle_d|^2/r_i) < \infty\}. \tag{17}$$

In $\mathcal{H}_n$ one has the symmetric relations.

Remark 6. If $|\chi\rangle_{nd}$ is infinitely entangled, then we have the following chain of ranges as proper subsets:

$$R(\rho_d) \subset R(\rho_d^{1/2}) \subset R(Q_d) \qquad (R(Q_d) = \bar{R}(\rho_d) = \bar{R}(\rho_d^{1/2}))$$

whereas, if the entanglement is finite, then all three linear manifolds coincide. Naturally, in $\mathcal{H}_n$ we have the symmetric relations.

Theorem 2. If a composite-system state vector $|\chi\rangle_{nd}$ is given, then a state vector $|\lambda\rangle_d$ can be obtained by distant preparation (cf proposition 1 and theorem 1) if and only if $|\lambda\rangle_d \in \mathrm{R}(\rho_d^{1/2})$.

Proof. As it is clear from proposition 1 and theorem 1, the (unnormalized) vector $|\lambda\rangle_d'$ is obtainable by distant preparation if and only if there exists a state vector $|\phi\rangle_n$ such that

$$\langle\phi|_n|\chi\rangle_{nd} = |\lambda\rangle_d'. \tag{18}$$

Replacing here (12), one obtains that the characteristic condition is existence of a $|\phi\rangle_n$ that satisfies

$$|\lambda\rangle_d' = \sum_i \langle\phi|_n|\phi_i\rangle_n r_i^{1/2}|\omega_i\rangle_d. \tag{19}$$

To prove the *necessity* of the claimed condition, let us assume that there exists a $|\phi\rangle_n$ satisfying (19). Its very existence makes $(\sum_i |\langle\phi_i|_n|\phi\rangle_n|^2) < \infty$ necessary. Further, (19) entails $\langle\phi|_n|\phi_i\rangle_n = \langle\omega_i|_d|\lambda\rangle_d'^2/r_i^{1/2}$. Substitution in the preceding inequality gives $[\sum_i (|\langle\omega_i|_d|\lambda\rangle_d'|^2/r_i)] < \infty$. On account of (17) this implies

$$|\lambda\rangle_d' \in R(\rho_d^{1/2}). \tag{20}$$

To prove *sufficiency*, we assume the validity of the claimed condition (20). Then, owing to (17), we can define

$$|\phi\rangle_n \equiv \sum_i (\langle\lambda|_d'|\omega_i\rangle_d/r_i^{1/2})|\phi_i\rangle_n$$

and, on account of (19), it determines by distant preparation the given $|\lambda\rangle_d'$. □

This result was derived in previous work (Herbut and Vujičić 1976) using antilinear Hilbert–Schmidt operator techniques. Therefore, the above proof may be viewed as a simplified one.

5. The central theorem on statistical teleportation

As it was stated, in the first step of statistical teleportation one measures $(|\bar{a}\rangle_{12}\langle\bar{a}|_{12} \otimes I_3)$ in the state $|\psi\rangle_1|\bar{\Psi}\rangle_{23}$. The nearby subsystem is $(1+2)$, and the distant one 3. Owing to theorem 2, we know that the distantly prepared state $|\psi\rangle'_3$ (cf the figure) necessarily belongs to the linear manifold $R((\bar{\rho}_3^B)^{1/2})$, where

$$\bar{\rho}_3^B \equiv \mathrm{Tr}_{12}(|\psi\rangle_1|\bar{\Psi}\rangle_{23}\langle\psi|_1\langle\bar{\Psi}|_{23}) = \mathrm{Tr}_2(|\bar{\Psi}\rangle_{23}\langle\bar{\Psi}|_{23}). \tag{21}$$

As it was stated, under the assumption that a bridge state vector $|\bar{\Psi}\rangle_{12}$ is given, we search for a state vector $|\bar{a}\rangle_{12}$ such that it makes possible the first step of statistical teleportation, i.e. distant preparation as a unitary map $\mathcal{U}_{31}$ of $\boldsymbol{S}_1$ onto $\boldsymbol{S}'_3$. The second step is then, in principle, defined as obvious from the figure:

$$\mathcal{U}_3 = \mathcal{I}_{31} \circ \mathcal{U}_{31}^{-1}.$$

We now fix *arbitrarily* $\boldsymbol{S}_1$ as a linear submanifold of $\mathcal{H}_1$ with finite or infinite dimensionality. We introduce the corresponding projector P_1 projecting onto the subspace $\bar{\boldsymbol{S}}_1$ (topological closure of $\boldsymbol{S}_1$).

Further, we define the linear submanifold $\boldsymbol{S}'_3$ of $\mathcal{H}_3$ as the one spanned by all distantly prepared state vectors $|\Psi\rangle'_3$ obtained from some vector $|\psi\rangle_1 \in \boldsymbol{S}_1$, and we denote by P'_3 the projector projecting $\mathcal{H}_3$ onto $\bar{\boldsymbol{S}}'_3$.

Definition 9. Let the *relevant bridge state vector* $|\Psi\rangle_{23}$ be defined by

$$|\Psi\rangle_{23} \equiv c^{-1/2}(I_2 \otimes P'_3)|\bar{\Psi}\rangle_{23} \tag{22}$$

where c is the square norm of the projection. The range projector of the state operator of the second subsystem in the relevant bridge state vector, i.e. of

$$\rho_2^B \equiv \mathrm{Tr}_3\,|\Psi\rangle_{23}\langle\Psi|_{23} \tag{23}$$

will be denoted by P_2.

To make sure that definition (22) is consistent, we must ascertain of the following fact.

Lemma 6. The projection $(I_2 \otimes P'_3)|\bar{\Psi}\rangle_{23}$ is never zero.

Proof. Assuming *ab contrario* that the projection is zero, and utilizing (21), we obtain

$$0 = \mathrm{Tr}_2(I_2 \otimes P'_3)|\bar{\Psi}\rangle_{23}\langle\bar{\Psi}|_{23}(I_2 \otimes P'_3) = P'_3\bar{\rho}_3^B P'_3.$$

(We have 'pulled out' P'_3 from Tr_2 (cf corollary 6)). Theorem 2 and the definition of $\boldsymbol{S}'_3$ imply $R(P'_3) \subseteq R[(\bar{\rho}_3^B)^{1/2}]$. Hence, if $P'_3|\phi\rangle_3 = |\phi\rangle_3$, then, on the one hand, $[(\bar{\rho}_3^B)^{1/2}]$ acts on $|\phi\rangle_3$ as a nonsingular operator, and, on the other hand, the above relation implies

$$0 = \langle\phi|_3(P'_3\bar{\rho}_3^B P'_3)|\phi\rangle_3 = \|((\bar{\rho}_3^B)^{1/2}|\phi\rangle_3\|^2$$

and, further, $(\bar{\rho}_3^B)^{1/2}|\phi\rangle_3 = 0$ in contradiction with the above. This completes our *reductio ad absurdum* proof. □

Definition 10. We call the state vector $|a\rangle_{12}$ defined by

$$|a>_{12} \equiv d^{-1/2}(P_1 \otimes P_2)|\bar{a}\rangle_{12} \tag{24}$$

where d is the square norm of the projection, *the relevant atom state vector*.

Lemma 7. The basic distant-preparation expression can be reduced to an analogous expression in terms of the relevant state vectors:

$$\forall|\psi\rangle_1 \in \boldsymbol{S}_1 : \langle\bar{a}|_{12}(|\psi\rangle_1|\bar{\Psi}\rangle_{23}) = c^{1/2}d^{1/2}\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}). \tag{25}$$

Proof. Utilizing first the fact that, by definition of $\boldsymbol{S}'_3$, all distantly prepared vectors belong to it, then corollary 4, then (22), then $|\psi\rangle_1 \in \boldsymbol{S}_1$ and the definition of P_2 together with corollary 10, then corollary 3, and finally (24), one can write:

$$\begin{aligned}\text{LHS} &= P'_3\langle\bar{a}|_{12}(|\psi\rangle_1|\bar{\Psi}\rangle_{23}) = \langle\bar{a}|_{12}\{|\psi\rangle_1[(I_2\otimes P'_3)|\bar{\Psi}\rangle_{23})]\}\\ &= c^{1/2}\langle\bar{a}|_{12}(|\psi\rangle_1|\Psi\rangle_{23}) = c^{1/2}\langle\bar{a}|_{12}[(P_1|\psi\rangle_1)(P_2\otimes I_3)|\Psi\rangle_{23}]\\ &= c^{1/2}[\langle\bar{a}|_{12}(P_1\otimes P_2)](|\psi\rangle_1|\Psi\rangle_{23}) = \text{RHS}.\end{aligned}$$

□

Proposition 2. If the measurement of the event $(|\bar{a}\rangle_{12}\langle\bar{a}|_{12}\otimes I_3)$ in the state $|\psi\rangle_1|\bar{\Psi}\rangle_{23}$ gives (by distant preparation) the state vector $|\psi\rangle'_3$ with probability $\bar{w}$, then the same state vector $|\psi\rangle'_3$ can be obtained with a probability w measuring the event $(|a\rangle_{12}\langle a|_{12}\otimes I_3)$ in the state $|\psi\rangle_1|\Psi\rangle_{23}$, where $|a\rangle_{12}$ is the relevant atom state vector of $|\bar{a}\rangle_{12}$ (cf (24)), and $|\Psi\rangle_{23}$ is the relevant bridge state vector (cf (22)). Besides, one has the relation:

$$\bar{w} = wcd. \tag{26}$$

Proof. Let us use the basic distant preparation formula (4) as the first step of teleportation in the two versions at issue:

$$\langle\bar{a}|_{12}(|\psi\rangle_1|\bar{\Psi}\rangle_{23}) \equiv (\bar{w})^{1/2}|\bar{\psi}\rangle'_3 \tag{27a}$$

$$\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}) = w^{1/2}|\psi\rangle'_3. \tag{27b}$$

As an immediate consequence of lemma 7, one can write

$$|\bar{\psi}\rangle'_3 = (\bar{w})^{-1/2}c^{1/2}d^{1/2}\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}) = (\bar{w})^{-1/2}c^{1/2}d^{1/2}w^{1/2}|\psi\rangle'_3.$$

Since we are dealing with state vectors and positive numbers, the claims $|\bar{\psi}\rangle'_3 = |\psi\rangle'_3$ and (26) follows. □

Remark 7. The distant-preparation relation (27*a*) is connected with the actual measurement (in the laboratory) of the event $(|\bar{a}\rangle_{12}\langle\bar{a}|_{12}\otimes I_3)$ in the state $|\psi\rangle_1|\bar{\Psi}\rangle_{23}$, i.e. with the first step of *actual* teleportation. Relation (27*b*), on the other hand, corresponds to *virtual* measurement of $(|a\rangle_{12}\langle a|_{12}\otimes I_3)$ in the state $|\psi\rangle_1|\Psi\rangle_{23}$. We call it 'virtual' because it is not actually performed, but it is conceivable.

We proceed in our search for atomic state vectors $|\bar{a}\rangle_{12}$ that will give teleportation. Further on, we utilize (27*b*) and virtual measurement, and we concentrate on its interpretation as a linear map $\mathcal{U}_{31}$.

Lemma 8. The relevant distant preparation relation (27*b*) amounts to a linear map $\mathcal{U}_{31}$ taking $\boldsymbol{S}_1$ into $\boldsymbol{S}'_3$ (cf (6*b*)) *if and only if* the probability w is one and the same positive number for all state vectors $|\psi\rangle_1 \in \boldsymbol{S}_1$.

Proof. For $|\psi\rangle_1, |\phi\rangle_1, |\omega\rangle_1 \in \boldsymbol{S}_1$, $\langle\psi|_1|\psi\rangle_1 = \langle\phi|_1|\phi\rangle_1 = \langle\omega|_1|\omega\rangle_1 = 1$, and $|\omega\rangle_1 = \alpha|\psi\rangle_1 + \beta|\phi\rangle_1, \alpha, \beta \in \boldsymbol{C}$, one can write

$$\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}) = w^{1/2}|\psi\rangle'_3 \tag{28a}$$

$$\langle a|_{12}(|\phi\rangle_1|\Psi\rangle_{23}) = x^{1/2}|\phi\rangle'_3 \qquad \langle a|(|\omega\rangle|\Psi\rangle) = y^{1/2}|\omega\rangle' \tag{28b}$$

and

$$\alpha\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}) + \beta\langle a|_{12}(|\phi\rangle_1|\Psi\rangle_{23}) = y^{1/2}|\omega\rangle'_3. \tag{29}$$

Sufficiency. We assume $w^{1/2} = x^{1/2} = y^{1/2} > 0$. Substituting on the LHS of (29) the equalities (28*a*), (28*b*), after cancellation one has

$$|\omega\rangle_3' = \alpha|\psi\rangle_3' + \beta|\phi\rangle_3' \tag{30}$$

as claimed.

Necessity. Let (30) be valid by assumption. Replacing (28*a*), (28*b*) and (30) in (29) one obtains

$$\alpha w^{1/2}|\psi\rangle_3' + \beta x^{1/2}|\phi\rangle_3' = \alpha y^{1/2}|\psi\rangle_3' + \beta y^{1/2}|\phi\rangle_3'.$$

If $|\psi\rangle_3'$ and $|\phi\rangle_3'$ are linearly independent, then the uniqueness of the expansion coefficients, and the specification $\alpha \neq 0 \neq \beta$ imply $w = x = y$. If we take $\alpha \neq 0 = \beta$ (or $\alpha = 0 \neq \beta$), then $|\omega\rangle_3'$, and $|\psi\rangle_3'$ (or $|\phi\rangle_3'$) are collinear, and it is clear from the distant-preparation relation (say (28*a*)) that the probability is the same. The positivity $w > 0$ is necessary, because otherwise $\mathcal{U}_{31}$ would map $\boldsymbol{S}_1$ into zero. □

We still do not know what kind of atomic state vector $|\bar{a}\rangle_{12}$ will give constant w over $\boldsymbol{S}_1$ in the relevant distant-preparation relation (27*b*). For further investigation, the Schmidt biorthogonal expansion (cf (12)) of the relevant bridge state vector $|\Psi\rangle_{23}$ is desirable. Let it be

$$|\Psi\rangle_{23} = \sum_i r_i^{1/2}|i\rangle_2|i\rangle_3. \tag{31a}$$

It is accompanied by the spectral forms of the subsystem state operators

$$\rho_2^B = \sum_i r_i|i\rangle_2\langle i|_2 \tag{31b}$$

$$\rho_3^B \equiv \mathrm{Tr}_2\,|\Psi\rangle_{23}\langle\Psi|_{23} = \sum_i r_i|i\rangle_3\langle i|_3 \tag{31c}$$

$\forall i : r_i > 0$ (cf (23) and (13*a*), (13*b*) *mutatis mutandis*).

Lemma 9. The relevant distant preparation relation (27*b*), i.e.

$$\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}) = w^{1/2}|\psi\rangle_3' \tag{32}$$

amounts to a unitary map $|\psi\rangle_3' = \mathcal{U}_{31}|\psi\rangle_1$ *if and only if* the following *suitability relation* is satisfied:

$$\rho_2 \equiv \mathrm{Tr}_1\,|a\rangle_{12}\langle a|_{12} = w(\tilde{\rho}_2^B)^{-1}P_2 \tag{33}$$

where $\tilde{\rho}_2^B$ is the restriction of ρ_2^B to its range $R(P_2)$. This implies

$$w = \{\mathrm{Tr}[(\tilde{\rho}_2^B)^{-1}]\}^{-1}. \tag{34}$$

Proof. It is useful to expand $|a\rangle_{12}$ in the ON subbasis $\{|i\rangle_2 : \forall i\}$ spanning $\boldsymbol{S}_2$ that is a characteristic basis of $\tilde{\rho}_2^B$:

$$|a\rangle_{12} = \sum_i (r_i^a)^{1/2}|i\rangle_1|i\rangle_2. \tag{35}$$

The subbasis $\{|i\rangle_1 : \forall i\}$ in $\boldsymbol{S}_1$ ($|a\rangle_{12} \in \boldsymbol{S}_1 \otimes \boldsymbol{S}_2$, cf (24)) *a priori* need not be orthogonal. We require the vectors $|i\rangle_1$ to be normalized, i.e. $\forall i : (r_i^a)^{1/2}$ is, by definition, the norm of the generalized expansion coefficient $(r_i^a)^{1/2}|i\rangle_1$ in (35) (cf lemma 2 *mutatis mutandis*).

Replacing expansions (35) and 31a) in the relevant distant-preparation relation (32), one obtains:

$$\sum_i \sum_{i'} [(r_i^a)^{1/2} r_{i'}^{1/2} \langle i|_1 |\psi\rangle_1 \langle i|_2 |i'\rangle_2] |i'\rangle_3 = w^{1/2} |\psi\rangle_3'.$$

Since $\langle i|_2 |i'\rangle_2 = \delta_{i,i}$, one ends up with

$$\sum_i (r_i^a)^{1/2} r_i^{1/2} \langle i|_1 |\psi\rangle_1 |i\rangle_3 = w^{1/2} |\psi\rangle_3'$$

or equivalently,

$$\forall i : w^{-1/2} (r_1^a)^{1/2} r_i^{1/2} \langle i|_1 |\psi\rangle_1 = \langle i|_3 |\psi\rangle_3'. \tag{36}$$

Since all steps that led from (32)–(36), made with the help of the expansions (31a) and (35), can be performed backwards, (36) is an equivalent form of (32). We must clarify under which conditions can (36) be viewed as a unitary isomorphism $\mathcal{U}_{31}$ mapping $\boldsymbol{S}_1$ onto $\boldsymbol{S}_3'$.

To prove *necessity* of the claimed condition (33), we assume that (36) determines a unitary isomorphism $\mathcal{U}_{31}$ ($\mathcal{U}_{31}|\psi\rangle_1 = |\psi\rangle_3'$) mapping $\boldsymbol{S}_1$ onto $\boldsymbol{S}_3'$.

Making $|\psi\rangle_1$ run over the ON basis vectors $\{|\bar{i}\rangle_1 \equiv \mathcal{U}_{31}^{-1} |i\rangle_3 : \forall i\}$, and taking into account the fact that w is one and the same by this (cf lemma 8), (36) implies

$$w^{-1/2} (r_i^a)^{1/2} r_i^{1/2} \langle i|_1 |\bar{i}'\rangle_1 = \delta_{i,i'}.$$

This makes the vectors $\{|i\rangle_1 : \forall i\}$, which were assumed to be normalized, also orthogonal on account of the orthonormality of the basis $\{|\bar{i}\rangle_1 : \forall i\}$, and it makes (35) a biorthogonal expansion. Besides, the last relation also implies

$$\forall i : r_i^a = w r_i^{-1}.$$

Since all r_i^a are seen to be positive, (35) is a Schmidt biorthogonal expansion (cf definition 4). Then, in view of the spectral form

$$\rho_2 = \sum_i r_i^a |i\rangle_1 \langle i|_1$$

which we can now read from (35) (cf corollary 8 *mutatis mutandis*), and taking into account (31b), the claimed relation (33) follows.

To prove *sufficiency* of the claimed condition (33), we point out that, due to the stated equivalence of (36) and (32), w in the two relations is one and the same, and hence it is the probability of distant preparation. Now, (33) implies (34), and hence the independence of w from the choice of $|\psi\rangle_1$ in $\boldsymbol{S}_1$. According to lemma 8, this is sufficient for the linearity (and hence the unitarity) of $\mathcal{U}_{31}$ determined by (32). □

Now we are prepared to state the central result of this study.

Theorem 3. Let $|\bar{\Psi}\rangle_{23}$ be a given (finitely or infinitely) entangled bridge state vector. Let the linear manifold $\boldsymbol{S}_1$ ($\subseteq \mathcal{H}_1$) be arbitrarily chosen but finitely dimensional, its number of dimensions not exceeding that of $R(\bar{\rho}_3^B)$. Then and only then there exist atomic state vectors $|\bar{a}\rangle_{12}$ such that the first step of statistical teleportation is possible.

A necessary and sufficient condition for the mentioned atomic state vector is that it defines a relevant atom state vector (i.e. that $d > 0$ in (24)), and that the latter, if written as a Schmidt biorthogonal expansion (35), satisfies:

(i) $\{|i\rangle_2 : \forall i\}$ is an arbitrary characteristic ON basis of ρ_2^B in $\boldsymbol{S}_2$;

(ii) $\forall i : r_i^a = w r_i^{-1}$ (cf (31a)–(31c)), and

(iii) w, the probability of statistical teleportation, is given by (34), i.e. by

$$w = \{\mathrm{Tr}[(\tilde{\rho}_2^B)^{-1}]\}^{-1}.$$

Proof. Proof follows immediately from lemma 9, because, in view of the spectral forms of ρ_2^B and ρ_2, (i) in conjunction with (ii) is equivalent to (33). The subspace $\boldsymbol{S}_1$ has to be finitely dimensional, because, in the first step of teleportation, $\boldsymbol{S}_1$, $\boldsymbol{S}_3'$ and $\boldsymbol{S}_2$ were seen to be equally dimensional (cf (6*b*) and (31*a*)), and for infinitely dimensional $\boldsymbol{S}_2$ the inversion in (33) would not be possible. Since $\boldsymbol{S}_3' \subseteq R(\bar{\rho}_3^B)$ (cf theorem 2 and remark 6), the number of dimensions of the former must not exceed that of the latter. These properties are seen to be also sufficient for (33), which, according to lemma 9, is sufficient for $|\bar{a}\rangle_{12}$ to be atomic if $d > 0$ in (24). □

Remark 8. Note that, while the spectrum $\{r_j^a : \forall j\}$ and the ON subbasis $\{|i\rangle_2 : \forall i\}$ in the Schmidt biorthogonal expansion (35) are determined by ρ_2^B, the ON subbasis $\{|i\rangle_1 : \forall i\}$ is an (arbitrary) ON basis in $\boldsymbol{S}_1$. Also the irrelevant parts of the atomic state vector $|\bar{a}\rangle_{12}$, i.e.

$$(P_1 \otimes P_2^{\perp})|\bar{a}\rangle_{12} \qquad (P_1^{\perp} \otimes P_2)|\bar{a}>_{12} \qquad \text{and} \qquad (P_1^{\perp} \otimes P_2^{\perp})|\bar{a}\rangle_{12}$$

($P_2^{\perp}$ being the orthocomplementary projector of P_2 in $\mathcal{H}_2$ etc) are completely arbitrary (cf (24)). This may prove useful in practice (which will hopefully be reached): it may help to find tractable atomic state vectors $|\bar{a}\rangle_{12}$.

Finally, the unitary isomorphism $\mathcal{U}_3$ (taking the subspace $\boldsymbol{S}_3'$ onto $\boldsymbol{S}_3$ ($\equiv \mathcal{I}_{31}\boldsymbol{S}_1$)), responsible for the second step of teleportation, is implied (as it was stated) as $\mathcal{U}_3 = \mathcal{I}_{31} \circ \mathcal{U}_{31}^{-1}$.

One should note that the larger the subspace $\boldsymbol{S}_1$ is, the more statistical teleportation can be accomplished with one and the same $|\bar{a}\rangle_{12}$ and one and the same $\mathcal{U}_3$. Also the probability $\bar{w}$ is one and the same.

Finally, it is important to point out that the central result is immediately applicable to deterministic teleportation. Namely, the probability w does not depend on the choice of the atom state. (It is implied solely by the bridge state.) Hence, it is valid simultaneously for all, what may be called, characteristic atomic state vectors of a deterministically teleporting observable (cf (9)).

The operator relation (33) is thus the searched for *necessary and sufficient condition* also for the observable in deterministic teleportation.

6. Immediate consequences of the central theorem

Corollary 11. If one denotes by $\mathcal{U}_a^B$ and $\mathcal{U}_a$ the correlation operator of the relevant bridge state vector $|\Psi\rangle_{23}$ and that of the relevant atom state vector $|a\rangle_{12}$ respectively (cf definition 5), then one has

$$\mathcal{U}_a = (\mathcal{U}_a^B)^{-1} \circ \mathcal{U}_{31} \tag{37}$$

(see the commuting lower triangle on figure 2 below).

Proof. The suitability relation (33) of the atomic state vector implies that each characteristic basis of ρ_2^B is also a characteristic basis of ρ_2. Let $\{|i\rangle_2 : \forall i\}$ be a common characteristic

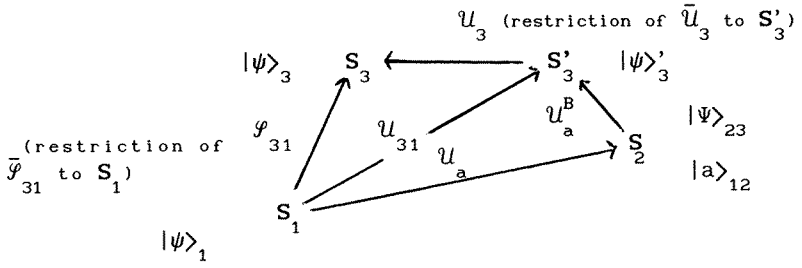


Figure 2.

orthonormal subbasis of ρ_2^B and ρ_2 spanning their common range $\boldsymbol{S}_2$. Then, utilizing (33), we can write down the Schmidt biorthogonal expansions as follows

$$|\Psi\rangle_{23} = \sum_{i=1}^{N} r_i^{1/2} |i\rangle_2 \otimes [(\mathcal{U}_a^B)|i\rangle_2]_3 \tag{38}$$

$$|a\rangle_{12} = \sum_{i=1}^{N} w^{1/2} r_i^{-1/2} [(\mathcal{U}_a)^{-1}|i\rangle_2]_1 \otimes |i\rangle_2 \tag{39}$$

(cf (15*a*) and (15*b*) *mutatis mutandis*).

By the very definition of $\mathcal{U}_{31}$ (and on account of proposition 2, stating that distant preparation can be written in terms of the relevant state vectors), one has:

$$\mathcal{U}_{31}|\psi\rangle_1 = |\psi\rangle'_3 = w^{-1/2}\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}).$$

Substitution of $|\psi\rangle_1 \equiv [(\mathcal{U}_a^{-1})|i\rangle_2]_1$, of (38), and of (39) entails

$$\forall i : \mathcal{U}_{31}[(\mathcal{U}_a^{-1})|i\rangle_2]_1 = [(\mathcal{U}_a^B)|i\rangle_2]_3.$$

On account of the commutativity of the triangle of unitary maps at issue (see figure 2), the claimed relation (37) then clearly follows. □

Corollary 12. In view of (34) and (38), one can write

$$w = 1/\sum_{i=1}^{N} r_i^{-1} \tag{40}$$

where N is the degree of entanglement of the relevant bridge state vector (cf definition 6 and (38)). It is always true that

$$w \leqslant 1/N^2 \tag{41}$$

and the maximal value $w = 1/N^2$ is reached if and only if the relevant bridge state vector is *maximally entangled*, i.e. $\forall i : r_i = 1/N$ (cf definition 7).

Proof. Relation (40) is the explicit form of (34). It is the harmonic mean value of the characteristic values r_i (of $\tilde{\rho}_2^B$) divided by N. Since the maximal value of the harmonic mean is the arithmetic mean, one obtains (41). The last claim follows from the fact that the maximal value of the harmonic mean, at fixed value of the arithmetic mean, is achieved if and only if the characteristic values are equal to each other. Since $\sum_i r_i = 1$, we are then dealing with a maximally entangled bridge state vector as claimed. □

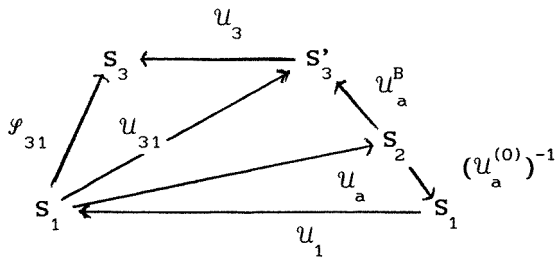


Figure 3.

Thus, from the statistical point of view it is most favourable to use a maximally entangled relevant bridge state vector for statistical teleportation because it gives the largest probability of teleportation.

Corollary 13. Let $|\bar{a}\rangle_{12}$ be an atomic state vector in $(\mathcal{H}_1 \otimes \mathcal{H}_2)$, i.e. a state vector such that (24) with $d > 0$ and (33) are valid. Let, further, $\mathcal{U}_a$ be the correlation operator of the implied atom state vector $|a\rangle_{12}$ mapping the range $R(\rho_1)$, where $\rho_1 \equiv \mathrm{Tr}_2 |a\rangle_{12}\langle a|_{12}$, onto $R(\rho_2)$ (cf (39)). Then, *the relevant part* $\mathcal{U}_3$ of the operation $\bar{\mathcal{U}}_3$, i.e. the map obtained by restricting the domain of the latter to $\boldsymbol{S}'_3$ $(\equiv R(\rho_3^B))$, is a unitary isomorphism mapping $\boldsymbol{S}'_3$ onto $\boldsymbol{S}_3$ $(\equiv \mathcal{I}_{31}R(\rho_1))$. The isomorphism $\mathcal{U}_3$ is uniquely determined by the two commuting triangles on figure 2 as follows

$$\mathcal{U}_3 = \mathcal{I}_{31} \circ \mathcal{U}_{31}^{-1} \qquad \mathcal{U}_{31} = \mathcal{U}_a^B \circ \mathcal{U}_a. \tag{42}$$

Proof. Obvious in view of the facts that the central theorem covers all possible atomic state vectors $|\bar{a}\rangle_{12}$ for fixed $\boldsymbol{S}_1$, and that the two basic ingredient entities ρ_2 and $\mathcal{U}_a$ of the relevant atom state vector $|a\rangle_{12}$ are decoupled (cf corollary 9). Note that figure 2, on which the mutual connections between the maps are displayed, is completely independent of ρ_2^B, which determines ρ_2. □

Corollary 14. Let $|\bar{a}^{(0)}\rangle_{12}$ be a fixed atomic state vector, and let $\mathcal{U}_a^{(0)}$ be the correlation operator of the implied relevant atom state vector $|a^{(0)}\rangle_{12}$ (cf definition 5). An arbitrary unitary operator $\mathcal{U}_1$ in $\boldsymbol{S}_1$ gives another relevant atom state vector $|a\rangle_{12}$ by specifying its correlation operator $\mathcal{U}_a$ as follows

$$\mathcal{U}_a^{-1} \equiv \mathcal{U}_1 \circ (\mathcal{U}_a^{(0)})^{-1} \tag{43}$$

(see figure 3) and ρ_2 is the same both for $|a^{(0)}\rangle_{12}$ and for $|a\rangle_{12}$, cf (33)). (The irrelevant parts of $|\bar{a}\rangle_{12}$, are, of course, arbitrary.) Every $|a\rangle_{12}$ can be obtained in this way.

Finally, the atom-generating unitary operators $\mathcal{U}_1$ and the operations $\mathcal{U}_3$ determine each other along the peripheral route on the commuting figure 3:

$$\mathcal{U}_3 = \mathcal{I}_{31} \circ \mathcal{U}_1 \circ (\mathcal{U}_a^{(0)})^{-1} \circ (\mathcal{U}_a^B)^{-1} \tag{44a}$$

$$\mathcal{U}_1 = \mathcal{I}_{31}^{-1} \circ \mathcal{U}_3 \circ \mathcal{U}_a^B \circ \mathcal{U}_a^{(0)}. \tag{44b}$$

Proof. Obvious. □

Corollary 15. If $|a^{(0)}\rangle_{12}$ is a given relevant atom state vector, and $\mathcal{U}_1$ is an arbitrary unitary operator in $\boldsymbol{S}_1$, then $|a\rangle_{12} \equiv (\mathcal{U}_1 \otimes I_2)|a^{(0)}\rangle_{12}$ is another relevant atom state vector, actually, the one described in the preceding corollary.

Proof. Follows immediately from (43) when one writes one relevant atom state vector in its Schmidt biorthogonal expansion, and one derives that of the other:

$$|a\rangle_{12} = \sum_i w^{1/2} r_i^{-1/2} (\mathcal{U}_a^{-1}|i\rangle_2)_1 |i\rangle_2 = \sum_i w^{1/2} r_i^{-1/2} \{(\mathcal{U}_1 \circ (\mathcal{U}_a^{(0)})^{-1})|i\rangle_2\}_1 |i\rangle_2$$
$$= (\mathcal{U}_1 \otimes I_2)|a^{(0)}\rangle_{12}.$$

□

7. Deterministic teleportation

We now return to the important concept of deterministic teleportation invented by Bennet *et al* (1993) (introduced in section 3). A complete observable A_{12} for subsystem $(1+2)$ is given such that all its characteristic vectors are atomic state vectors (see relation (9)):

$$A_{12} = \sum_{m=1}^{N} a_m |\bar{a}^{(m)}\rangle_{12} \langle \bar{a}^{(m)}|_{12} \tag{45a}$$

with $m \neq m' \Rightarrow a_m \neq a_{m'}$, and

$$\langle \bar{a}^{(m)}|_{12} |\bar{a}^{(m')}\rangle_{12} = \delta_{m,m'}. \tag{45b}$$

Since A_{12} is a complete observable in $(\mathcal{H}_1 \otimes \mathcal{H}_2)$ ($\mathcal{H}_i$ being the state space of subsystem i, $i = 1, 2$), the set $\{|\bar{a}^{(m)}\rangle_{12} : m = 1, 2, \ldots, M\}$ must span this space. Hence, M must equal the product of the numbers of dimensions of $\mathcal{H}_1$ and of $\mathcal{H}_2$ (finite or infinite).

Theorem 4. Deterministic teleportation with an entangled bridge state vector $|\bar{\Psi}\rangle_{23}$ and with a finitely-dimensional subspace $\boldsymbol{S}_1$ is possible *only if* the following condition is satisfied:

$$\sum_{m=1}^{M} d_m = w^{-1} c^{-1} \tag{46}$$

where w is given by (34), c is given by (22), and d_m is determined by

$$d_m^{1/2} |a^{(m)}\rangle_{12} \equiv (P_1 \otimes P_2)|\bar{a}^{(m)}\rangle_{12} \qquad m = 1, 2, \ldots, M \tag{47}$$

(cf (24)).

Proof. One has deterministic teleportation if and only if the observable A_{12} given in spectral form by (45*a*) is complete and each of its characteristic vectors is an atomic state vector. Then one of the atomic events $(|\bar{a}^{(m)}\rangle_{12} \langle \bar{a}^{(m)}|_{12} \otimes I_3)$ necessarily does occur in the measurement of the observable in the state $|\psi\rangle_1 |\bar{\Psi}\rangle_{23}$. Hence, the corresponding probabilities $\bar{w}_m$ $(= wcd_m$, cf (26)) of the mentioned events must add up into 1. This can obviously be written as (46). □

Remark 9. If there is no redundancy in the characteristic atomic state vectors $|\bar{a}^{(m)}\rangle_{12}$ of a deterministically teleporting observable A_{12} (cf (45*a*), (45*b*)), i.e. if $|\bar{a}^{(m)}\rangle_{12} = |a^{(m)}\rangle_{12}$, $m = 1, 2, \ldots, M$ (cf (24)), then, on the one hand, each of the vectors belongs to $(\boldsymbol{S}_1 \otimes \boldsymbol{S}_2)$, and on the other hand, A_{12} must be complete in $(\mathcal{H}_1 \otimes \mathcal{H}_2)$, or, equivalently the basis $\{|a^{(m)}\rangle_{12} : m = 1, 2, \ldots, M\}$ must span the entirety of this space. Hence, necessarily $\boldsymbol{S}_1 = \mathcal{H}_1$, and $\boldsymbol{S}_2 = \mathcal{H}_2$, and $\mathcal{H}_k$, $k = 1, 2$, must be *finite dimensional*. They must also be *equally dimensional* because the number of dimensions of both must equal that of $\boldsymbol{S}_3'$ (on account of $\mathcal{U}_{31}$ and the definition of $\boldsymbol{S}_2$).

If there is redundancy, then, it appears, there is the possibility of using any entangled bridge state vector $|\bar{\Psi}\rangle_{23}$, even an infinitely entangled one, for deterministic teleportation, as well as an infinitely dimensional state space $\mathcal{H}_1$. Namely, however large the RHS of (46), the numerous small d_m might add up into it.

Theorem 5. In the special case when $\boldsymbol{S}_1 = \mathcal{H}_1$ and $\boldsymbol{S}_2 = \mathcal{H}_2$, deterministic teleportation is possible if and only if there is no redundancy either in the bridge state vector or in any characteristic state vector, and if the former is *maximally entangled*.

Proof. Let N be the degree of entanglement of $|\Psi\rangle_{23}$ (and, of course, the number of dimensions of $\boldsymbol{S}_i, i = 1, 2$). Relation (46) reads $\sum_{m=1}^{M} w d_m c = 1$, and $M = N^2$. Since

$$w \leqslant 1/N^2 \qquad c \leqslant 1 \qquad d_m \leqslant 1 \qquad m = 1, 2, \ldots, N^2$$

(cf (41), (22), and (47)), relation (46) can be satisfied if and only if each term has its maximal value $(1/N^2)$, i.e.

$$w = 1/N^2 \qquad c = d_m = 1 \qquad m = 1, 2, \ldots, N^2$$

and the first of these conditions means maximal entanglement (cf corollary 12). □

Corollary 16. Let again $\boldsymbol{S}_1 = \mathcal{H}_1, \boldsymbol{S}_2 = \mathcal{H}_2$. Let, further, $\{|i\rangle_2 : i = 1, 2, \ldots, N\}$ be a characteristic orthonormal basis of $\tilde{\rho}_2^B$ $(= \rho_2^B$, cf (23)). It is a characteristic orthonormal basis also of ρ_2 of each $|\bar{a}^{(m)}\rangle_{12}$ on account of (33). Hence, expansion of each $|\bar{a}^{(m)}\rangle_{12}$ in this basis is a Schmidt biorthogonal expansion (cf lemma 4 and definition 4):

$$|\bar{a}^{(m)}\rangle_{12} = |a^{(m)}\rangle_{12} = N^{-1/2} \sum_{i=1}^{N} |i^{(m)}\rangle_1 |i\rangle_2 \qquad m = 1, 2, \ldots, N^2 \tag{48}$$

(cf theorem 5). The *orthogonality* (45*b*) is equivalent to the $N(N-1)/2$ conditions

$$\forall (m < m') : \qquad \sum_{i=1}^{N} \langle i^{(m)}|_1 |i^{(m')}\rangle_1 = 0. \tag{49}$$

Proof. Immediately obtained when one replaces (48) in (45*b*). □

Corollary 17. If one has one observable A_{12} suitable for deterministic teleportation, any unitary operator $\mathcal{U}_1$ and $\mathcal{H}_1$ generates another:

$$A'_{12} \equiv (\mathcal{U}_1 \otimes I_2) A_{12} (\mathcal{U}_1^{\dagger} \otimes I_2).$$

Proof. This is an immediate consequence of (48) and (49). □

Corollary 18. If, under the conditions of theorem 5, one fixes one of the characteristic atom state vectors, e.g. $|a^{(m_0)}\rangle_{12}$, and one wants to generate the rest of the $(N^2 - 1)$ characteristic atom state vectors of a deterministically teleporting observable A_{12} out of it using unitary operators $\mathcal{U}_1^{(m)}$ (cf corollary 5), then *the orthogonality relations* (49) take on the *equivalent* form

$$\mathrm{Tr}(\mathcal{U}_1^{(m)})^{\dagger} \mathcal{U}_1^{(m')} = 0 \qquad m \neq m' \qquad m, m' = 1, 2, \ldots, N^2. \tag{50}$$

Since $\mathcal{U}_1^{(m_0)} = I_1$, (50) is seen to imply

$$\mathrm{Tr}\,\mathcal{U}_1^{(m)} = 0 \qquad \forall m, m \neq m_0. \tag{51}$$

Proof. Follows immediately from (49). □

Evidently, in our search for suitable pairs

$$\{(|a^{(m)}\rangle_{12}, \mathcal{U}_3^{(m)}), m = 1, 2, \ldots, N^2\}$$

for deterministic teleportation one can choose the unitary operators $\mathcal{U}_3^{(m)}$ as the *starting entities* and evaluate $(\mathcal{U}_a^{(m)})$ from (42). Then, under the conditions of theorem 5, the orthogonality relations (49) can be further transformed. Since, moving on the periphery of figure 3, one can write

$$\mathcal{U}_3^{(m)} = \mathcal{I}_{31} \circ \mathcal{U}_1^{(m)} \circ (\mathcal{U}_a^{(m_0)})^{-1} \circ (\mathcal{U}_a^B)^{-1} \qquad m = 1, 2, \ldots, N^2 \tag{52}$$

and since all the factors are unitary or antiunitary isomorphisms, which implies that adjoining is the same as inversion, one obtains

$$\begin{aligned} &[(\mathcal{U}_3^{(m)})^\dagger \mathcal{U}_3^{(m')}] = (\mathcal{U}_a^B \circ \mathcal{U}_a^{(m_0)}) \circ [(\mathcal{U}_1^{(m)})^\dagger \mathcal{U}_1^{(m')}] \circ (\mathcal{U}_a^B \circ \mathcal{U}_a^{(m_0)})^{-1} \\ &m \neq m' \qquad m, m' = 1, 2, \ldots, N^2. \end{aligned} \tag{53}$$

Thus, the square-bracketed operators are unitary transforms of each other, hence, their traces are equal. This proves the following result.

Corollary 19. Under the conditions of theorem 5, the orthogonality relations (45*b*) are equivalent to

$$\mathrm{Tr}[(\mathcal{U}_3^{(m)})^\dagger \mathcal{U}_3^{(m')}] = 0 \qquad m \neq m' \qquad m, m' = 1, 2, \ldots, N^2. \tag{54}$$

Remark 10. If one solves the orthogonality relations (54), then one derives the ON basis $\{|i^{(m)}\rangle_1 : \forall i\}$ that gives $|a^{(m)}\rangle_{12}$ via (48) utilizing (52) solved for $\mathcal{U}_1^{(m)}$:

$$\forall m, \forall i: \qquad |i^{(m)}\rangle_1 \equiv [(\mathcal{I}_{31}^{-1}) \circ (\mathcal{U}_3^{(m)}) \circ (\mathcal{U}_a^B) \circ (\mathcal{U}_a^{(m_0)})]|\mathrm{i}^{(m_0)} \circ \rangle_1. \tag{55}$$

Remark 11. Let us, for a moment, consider statistical teleportation by itself (not as a constituent of deterministic teleportation). Then, one must realize that if the event determined by the atomic state vector does not occur in its measurement, the state of subsystem 1 changes, and the same individual sample of subsystem 1 is not available for a second attempt. If the measurement is ideal (cf Messiah 1961, Lüders 1951), one can work out this change of state. (But it is not worth doing.) If the measurement is not ideal (as most realistic cases are), then even in principle one does not know the change of state.

Remark 12. It is important to bear in mind that for distant preparation (and for teleportation obtained after the operation $\mathcal{U}_3$ is applied) the measurement of the deterministically teleporting observable *need not be an ideal one* (cf section 6(B) in Herbut and Vujičić 1976).

Remark 13. As an example of derivation of deterministically teleporting observables A_{12} under the conditions of theorem 5, we take $N = 2$, and give a family of ON bases $\{\{|i^{(m)}\rangle_1 : i = 1, 2\} : m = 1, 2, 3, 4\}$ (which then, using (48) and satisfying (45*b*) via (49), each define an observable (45*a*), where the corresponding operations $\mathcal{U}_3^{(m)}$ are given by (52)):

$$\begin{aligned} &\{|_\uparrow\rangle_1, |^\downarrow\rangle_1\} \qquad \{\mathrm{e}^{\mathrm{i}\theta}|^\downarrow\rangle_1, \mathrm{e}^{-\mathrm{i}\theta}|_\uparrow\rangle_1\}. \\ &\{((1-s^2)^{1/2}|_\uparrow\rangle_1 + \mathrm{i}s\mathrm{e}^{\mathrm{i}\theta}|^\downarrow\rangle_1), (-\mathrm{i}s\mathrm{e}^{-\mathrm{i}\theta}|_\uparrow\rangle_1 - (1-s^2)^{1/2}|^\downarrow\rangle_1)\} \\ &\{(-\mathrm{i}s|_\uparrow\rangle_1 - (1-s^2)^{1/2}\mathrm{e}^{\mathrm{i}\theta}|^\downarrow\rangle_1), ((1-s^2)^{1/2}\mathrm{e}^{-\mathrm{i}\theta}|_\uparrow\rangle_1 + \mathrm{i}s|^\downarrow\rangle_1)\} \end{aligned} \qquad 0 \leqslant s < 1.$$

For $\theta = 0$ and $s = 0$ we have the special case of Bell's observable utilized by Bennett *et al* (1993).

The general solution for $N = 2$ requires a more complicated presentation. Therefore we omit it in this illustration.

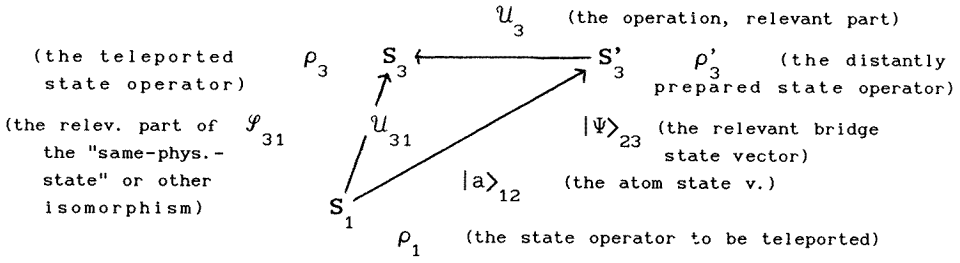


Figure 4.

8. Teleportation of proper mixtures

First, we generalize *statistical teleportation* to *general* (mixed or pure) *states*. We shall deal with state operators (statistical operators) instead of state vectors in this case.

One expects that generalizing the state vectors $|\psi\rangle_1$, $|\psi\rangle_3'$ $(= \mathcal{U}_{31}|\psi\rangle_1)$, and $|\psi\rangle_3$ $(= (\mathcal{U}_3 \circ \mathcal{U}_{31})|\psi\rangle_1 = \mathcal{I}_{31}|\psi\rangle_1)$ to state operators ρ_1, ρ_3' and ρ_3 via the corresponding similarity transformations

$$\rho_3' \ (= \mathcal{U}_{31}\rho_1\mathcal{U}_{31}^{-1}) \qquad \rho_3 \ (= (\mathcal{U}_3 \circ \mathcal{U}_{31})\rho_1(\mathcal{U}_3 \circ \mathcal{U}_{31})^{-1} = \mathcal{I}_{31}\rho_1\mathcal{I}_{31}^{-1})$$

one achieves statistical teleportation of the state operator.

We want to achieve figure 4.

One wonders what the counterpart of the restriction $|\psi\rangle_1 \in \boldsymbol{S}_1$ for the state operator ρ_1 is. The answer to this question is important because all that was stated about ρ_1 above cannot be expected to be valid in more generality than under the sought for restriction. We shall say that ρ_1 satisfying this restriction is 'suitable'.

Definition 11. A state operator ρ_1 will be called *suitable* if it satisfies one of the following four equivalent conditions.

(i) (*Viewing ρ_1 as one mixture.*) There exists a decomposition of ρ_1 into pure states, $\rho_1 = \sum_k w_k|\psi_k\rangle_1\langle\psi_k|_1$, such that $\forall k : |\psi_k\rangle_1 \in \boldsymbol{S}_1$.

(ii) (*Algebraic characterization.*) $P_1\rho_1 = \rho_1$.

(iii) (*Geometric characterization.*) The subspace $\boldsymbol{S}_1$ is invariant for ρ_1 and the latter reduces into zero in the orthocomplement $\boldsymbol{S}_1^{\perp}$.

(iv) (*Viewing ρ_1 as an arbitrary mixture of pure states.*) For every decomposition $\rho_1 = \sum_k w_k|\psi_k\rangle_1\langle\psi_k|_1$ of ρ_1 into pure states one has $\forall k : |\psi_k\rangle_1 \in \boldsymbol{S}_1$.

It will be proved in appendix A that the four restrictions on ρ_1 displayed in definition 11 are equivalent.

A plausibility argument for statistical teleportation of state operators goes as follows. Thinking of ρ_1 as of a mixture of pure states, e.g.

$$\rho_1 = \sum_k w_k|\psi_k\rangle_1\langle\psi_k|_1$$

($\forall k : w_k > 0$, $\langle\psi_k|_1|\psi_k\rangle_1 = 1$, and $|\psi_k\rangle_1 \in \boldsymbol{S}_1$, $\sum_k w_k = 1$), one imagines that an individual first subsystem is necessarily in one of the states $|\psi_k\rangle_1$, and it is thus teleported with the probability $\bar{w}$ (cf (26)) that does not depend either on k or w_k (cf (34)). This uniformity of $\bar{w}$ with respect to the above decomposition suggests that ρ_1 is teleported as a whole.

However, whenever ρ_1 is a mixed state, the above decomposition is necessarily nonunique and imagining ρ_1 as an actual physical mixture of pure states is open to doubt in quantum mechanics. Hence, it is desirable to have a rigorous argument that applies to

statistical operators as *compact entities*. This is the usual procedure in quantum mechanics, and we proceed along these lines also for teleportation.

The first step, i.e. *distant preparation*, is crucial in statistical teleportation. To begin with, we generalize the most important results of distant-preparation theory (see section 2) to statistical operators. We return to the notation used in section 2.

Proposition 3. Let ρ_{nd} be a given general composite-system state operator (in $(\mathcal{H}_n \otimes \mathcal{H}_d)$), and $|b\rangle_n$ a given nearby-subsystem state vector. Then ideal measurement of the event $(|b\rangle_n\langle b|_n \otimes I_d)$ in the state ρ_{nd}, if the event does occur, brings about distant preparation of the distant subsystem in the state

$$\rho'_d \equiv p^{-1}\,\mathrm{Tr}_n\,\rho_{nd}(|b\rangle_n\langle b|_n \otimes I_d) \tag{56a}$$

where p is the probability of occurrence

$$p \equiv \mathrm{Tr}_{nd}\,\rho_{nd}(|b\rangle_n\langle b|_n \otimes I_d) = \langle b|_n\rho_n|b\rangle_n \qquad \rho_n \equiv \mathrm{Tr}_d\,\rho_{nd}. \tag{56b}$$

Proof. Ideal occurrence brings about a change of state given by the Lüders formula (Lüders 1951, Messiah 1961):

$$\rho_{nd} \to \rho'_{nd} \equiv p^{-1}(|b\rangle_n\langle b|_n \otimes I_d)\rho_{nd}(|b\rangle_n\langle b|_n \otimes I_d).$$

One has

$$\mathrm{Tr}_n[(E_n \otimes I_d)B_{nd}] = \mathrm{Tr}_n[B_{nd}(E_n \otimes I_d)]$$

where E_n is a projector in $\mathcal{H}_n$ and B_{nd} is an everywhere defined linear operator in $\mathcal{H}_{nd}$. The equality can be seen to hold if the partial trace is evaluated in a characteristic basis of E_n. Since $\rho'_d \equiv \mathrm{Tr}_n\,\rho'_{nd}$, the above commutation under the partial trace and the projector idempotency lead to (56*a*). □

Theorem 6. Any direct measurement of an event $(|b\rangle_n\langle b|_n \otimes I_d)$ in the state ρ_{nd} results in the distant preparation of the state operator ρ'_d given by (56*a*) as in the case of ideal measurement (and, of course, the probability p given by (56*b*) is the same).

Proof. Obtainable by direct generalization of that of theorem 1, in particular, by replacing $\langle\chi|_{nd}\ldots|\chi\rangle_{nd}$ in (5) by $(\mathrm{Tr}_{nd}\,\rho_{nd}\ldots)$. □

We now return to our investigation of teleportation.

Proposition 4. Let

$$\rho_1 = \sum_k w_k|\psi_k\rangle_1\langle\psi_k|_1 \tag{57a}$$

be a given decomposition into pure states of a state operator ρ_1 satisfying restriction (i) in definition 11. Let, further, $|\bar{a}\rangle_{12}$ be an atomic state vector in $(\mathcal{H}_1 \otimes \mathcal{H}_2)$. Let, finally, a measurement of $(|\bar{a}\rangle_n\langle\bar{a}|_n \otimes I_d)$ in the three-subsystem composite-system state $(|\psi_k\rangle_1 \otimes |\bar{\Psi}\rangle_{23})$, if the event occurs, give by distant preparation the third-subsystem state $|\psi_k\rangle'_3$ for each value of k. Then, the same measurement in the state $(\rho_1 \otimes |\bar{\Psi}\rangle_{23}\langle\bar{\Psi}|_{23})$ in case of occurrence, gives by distant preparation the state ρ'_3 satisfying

$$\rho'_3 = \sum_k w_k|\psi_k\rangle'_3\langle\psi_k|'_3. \tag{57b}$$

Proof. According to theorem 6, one has

$$\rho_3' = (\bar{w})^{-1}\,\mathrm{Tr}_{12}(\rho_1 \otimes |\bar{\Psi}\rangle_{23}\langle\bar{\Psi}|_{23})(|\bar{a}\rangle_{12}\langle\bar{a}|_{12} \otimes I_3).$$

Substitution of decomposition (57*a*) gives further

$$\rho_3' = \sum_k w_k(\bar{w})^{-1}\langle\bar{a}|_{12}|\psi_k\rangle_1|\bar{\Psi}\rangle_{23}\langle\psi_k|_1\langle\bar{\Psi}|_{23}|\bar{a}\rangle_{12} = \sum_k w_k|\psi_k\rangle_3'\langle\psi_k|_3'$$

(cf theorem 1 and relation (4), as well as proposition 2). □

Theorem 7. If, under the assumptions of proposition 4, ρ_3' is the state operator to which ρ_1 gives rise by distant preparation, then one can write

$$\rho_3' = \mathcal{U}_{31}\rho_1\mathcal{U}_{31}^{-1}. \tag{58}$$

Proof. Follows immediately if one substitutes

$$\forall k : |\psi_k\rangle_3' = \mathcal{U}_{31}|\psi_k\rangle_1$$

in (57*b*) and one takes into account (57*a*). □

The two most important consequences are now obvious.

Corollary 20. Under the assumptions given in proposition 4, one has

$$\rho_3 \equiv (\mathcal{U}_3 \circ \mathcal{U}_{31})\rho_1(\mathcal{U}_3 \circ \mathcal{U}_{31})^{-1} = \mathcal{I}_{31}\rho_1\mathcal{I}_{31}^{-1}. \tag{59}$$

Corollary 21. If A_{12} is a complete observable defined in terms of atomic characteristic state vectors throughout like the one given by (45*a*), (45*b*), then its measurement in the state $(\rho_1 \otimes |\bar{\Psi}\rangle_{23}\langle\bar{\Psi}|_{23})$, where ρ_1 is a suitable state operator (cf definition 11), necessarily results in teleportation expressed by (59), where $\mathcal{U}_3$ and $\mathcal{U}_{31}$ have to be replaced by $\mathcal{U}_3^{(m)}$ and $\mathcal{U}_{31}^{(m)}$ respectively, m being the index of the atomic event $(|\bar{a}_{12}^{(m)}\rangle\langle\bar{a}|_{12}^{(m)} \otimes I_3)$ that did occur for the given individual system.

9. Teleportation of improper mixtures

We now assume that one more subsystem has entered the scene. We denote it by zero. Let ρ_{01} be an arbitrary state operator of the composite system $(0+1)$. The state operator ρ_1 $(\equiv \mathrm{Tr}_0\,\rho_{01})$ of subsystem 1 describes now an *improper mixture* (cf D'Espagnat 1976) unless ρ_{01} is uncorrelated (i.e. unless $\rho_{01} = \rho_0 \otimes \rho_1$, where $\rho_0 \equiv \mathrm{Tr}_1\,\rho_{01}$).

The *correlations* in a correlated state operator ρ_{01} are given in terms of the totality of *conditional state operators* $\{\rho_0(F_1)$: all $F_1\}$ acting in $\mathcal{H}_0$, where an arbitrary event (projector) F_1 for subsystem 1 is the condition. The state $\rho_0(F_1)$ of subsystem 0 comes about when $(I_0 \otimes F_1)$ (I_0 being the identity operator in $\mathcal{H}_0$) is measured in the state ρ_{01} and the event occurs (for more details see section 2 in Herbut 1986).

Theorem 8. Let

$$|\chi\rangle_{01} = \sum_i r_i^{1/2}|b_i\rangle_0|c_i\rangle_1 \tag{60}$$

be a state vector in $(\mathcal{H}_0 \otimes \mathcal{H}_1)$ given in a Schmidt biorthogonal expansion (cf definition 4) and such that the state operator of subsystem 1, i.e. ρ_1 $(\equiv \mathrm{Tr}_0\,|\chi\rangle_{01}\langle\chi|_{01})$, satisfies the suitability restriction (cf definition 11). Let, further, $|\bar{\Psi}\rangle_{23}$ be a given bridge state vector, and let $|\bar{a}\rangle_{12}$ be an atomic state vector. Finally, let the event $(I_0 \otimes |\bar{a}\rangle_{12}\langle\bar{a}|_{12} \otimes I_3)$ be

measured in the state $|\chi\rangle_{01}|\bar{\Psi}\rangle_{23}$. Then, if the event occurs, *distant preparation*, i.e. the first step of statistical teleportation, brings about a state vector $|\chi\rangle'_{03}$ so that

$$|\chi\rangle'_{03} = (I_0 \otimes \mathcal{U}_{31})|\chi\rangle_{01} \tag{61}$$

where $\mathcal{U}_{31}$ is the unitary isomorphism mapping $\boldsymbol{S}_1$ onto $\boldsymbol{S}'_3$ that is determined by $|\bar{a}\rangle_{12}$ and $|\bar{\Psi}\rangle_{23}$ (cf (6*a*) and (6*b*)).

Proof. Utilizing the basic distant-preparation formula (4) (cf also Theorem 1), and taking subsystem $(1+2)$ as the nearby and $(0+3)$ as the distant subsystem, one has

$$\begin{aligned}|\chi\rangle'_{03} &= (\bar{w})^{-1/2}\langle\bar{a}|_{12}\left(\sum_i r_i^{1/2}|b_i\rangle_0|c_i\rangle_1\right)|\bar{\Psi}\rangle_{23} \\ &= \sum_i r_i^{1/2}|b_i\rangle_0[(\bar{w})^{-1/2}\langle\bar{a}|_{12}|c_i\rangle_1)|\bar{\Psi}\rangle_{23}] = \sum_i r_i^{1/2}|b_i\rangle_0(\mathcal{U}_{31}|c_i\rangle_1) \\ &= (I_0 \otimes \mathcal{U}_{31})\left(\sum_i r_i^{1/2}|b_i\rangle_0|c_i\rangle_1\right) = (I_0 \otimes \mathcal{U}_{31})|\chi\rangle_{01}.\end{aligned}$$

Use has been made of (6*a*) and (6*b*) and of the fact that $\bar{w}$ does not depend on the state vector from $\boldsymbol{S}_1$. □

Corollary 22. Under the assumptions of theorem 8, *statistical teleportation* gives

$$|\chi\rangle_{03} = (I_0 \otimes \mathcal{U}_3) \circ (I_0 \otimes \mathcal{U}_{31})|\chi\rangle_{01} = (I_0 \otimes \mathcal{I}_{31})|\chi\rangle_{01}. \tag{62}$$

Proof. Obvious. □

Corollary 23. The *correlations* in a teleported state vector $|\chi\rangle_{01}$ are *preserved* in the sense that we have the Schmidt biorthogonal expansions

$$|\chi\rangle'_{03} = (I_0 \otimes \mathcal{U}_{31})|\chi\rangle_{01} = \sum_i r_i^{1/2}|b_i\rangle_0(\mathcal{U}_{31}|c_i\rangle_1)_3 \tag{63}$$

$$|\chi\rangle_{03} = (I_0 \otimes \mathcal{I}_{31})|\chi\rangle_{01} = \sum_i r_i^{1/2}|b_i\rangle_0(\mathcal{I}_{31}|c_i\rangle_1)_3 \tag{64}$$

essentially unchanged with respect to (60), i.e. unchanged in $\mathcal{H}_0$.

Proof. Follows immediately from (61) and (62). □

Theorem 9. Let $|\bar{\Psi}\rangle_{23}$ be a given bridge state vector, let $|\bar{a}\rangle_{12}$ be an atomic state vector, and let the event $(I_0 \otimes |\bar{a}\rangle_{12}\langle\bar{a}|_{12} \otimes I_3)$ be measured in the state given by the state operator $\rho_{01} \otimes (|\bar{\Psi}\rangle_{23}\langle\bar{\Psi}|_{23})$, where ρ_{01} is an arbitrary state operator in $(\mathcal{H}_0 \otimes \mathcal{H}_1)$, but such that the state operator of subsystem 1, i.e. ρ_1 $(\equiv \mathrm{Tr}_0\,\rho_{01})$, satisfies the suitability restriction (definition 11). Then, in case of occurrence, the distantly prepared state operator ρ'_{03} is obtained from ρ_{01} as follows

$$\rho'_{03} = (I_0 \otimes \mathcal{U}_{31})\rho_{01}(I_0 \otimes \mathcal{U}_{31})^{-1} \tag{65}$$

where $\mathcal{U}_{31}$ is the unitary isomorphism mapping $\boldsymbol{S}_1$ onto $\boldsymbol{S}'_3$ that is determined by $|\bar{a}\rangle_{12}$ and $|\bar{\Psi}\rangle_{23}$ (cf (6*a*) and (6*b*)).

Proof. Taking the subsystem $(1+2)$ as the nearby and $(0+3)$ as the distant subsystem, according to theorem 6 distant preparation gives

$$\begin{aligned}\rho'_{03} &= (\bar{w})^{-1}\,\mathrm{Tr}_{12}\{[\rho_{01}\otimes(|\bar{\Psi}\rangle_{23}\langle\bar{\Psi}|_{23})][I_0\otimes|\bar{a}\rangle_{12}\langle\bar{a}|_{12}\otimes I_3]\}\\ \bar{w} &= \mathrm{Tr}_{0123}\{[\rho_{01}\otimes(|\bar{\Psi}\rangle_{23}\langle\bar{\Psi}|_{23})][I_0\otimes|\bar{a}\rangle_{12}\langle\bar{a}|_{12}\otimes I_3]\}.\end{aligned} \tag{66}$$

Let

$$\rho_{01}=\sum_k w_k|\chi_k\rangle_{01}\langle\chi_k|_{01} \tag{67}$$

be a decomposition of ρ_{01} into pure states. Replacing this in (66), using the fact that $\bar{w}$ does not depend on $|\chi_k\rangle_{01}$, and utilizing (61), we derive

$$\rho'_{03}=\sum_k w_k(I_0\otimes\mathcal{U}_{31})|\chi_k\rangle_{01}\langle\chi_k|_{01}(I_0\otimes\mathcal{U}_{31})^{-1}$$

which, on account of (67), implies (65). □

Corollary 24. Under the assumptions of theorem 9, the statistically teleported state operator can be written

$$\rho_{03}=[(I_0\otimes\mathcal{U}_3)\circ(I_0\otimes\mathcal{U}_{31})]\rho_{01}[(I_0\otimes\mathcal{U}_3)\circ(I_0\otimes\mathcal{U}_{31})]^{-1}=(I_0\otimes I_{31})\rho_{01}(I_0\otimes I_{31})^{-1}. \tag{68}$$

Proof. Obvious. □

Theorem 10. Let F_1 be an arbitrary subprojector of P_1, i.e. such that $F_1P_1=F_1$ is valid (in view of P_1, cf definition 11). Then *the conditional state operator* in $\mathcal{H}_0$, i.e.

$$\rho_0(F_1)\equiv p^{-1}\,\mathrm{Tr}_1[\rho_{01}(I_0\otimes F_1)] \tag{69a}$$

where

$$p\equiv\mathrm{Tr}_{01}[\rho_{01}(I_0\otimes F_1)] \tag{69b}$$

is *preserved* in the sense that

$$\rho_0(F_1)=\rho_0(\mathcal{U}_{31}F_1\mathcal{U}_{31}^{-1}) \tag{70a}$$

$$=\rho_0(\mathcal{I}_{31}F_1\mathcal{I}_{31}^{-1}). \tag{70b}$$

Proof. Denoting by p' the probability that $[I_0\otimes(\mathcal{U}_{31}F_1\mathcal{U}_{31}^{-1})]$ occurs in the state $\rho'_{03'}$ and utilizing (65), one has

$$\begin{aligned}\rho_0(\mathcal{U}_{31}F_1\mathcal{U}_{31}^{-1}) &\equiv (p')^{-1}\mathrm{Tr}_3\{\rho'_{03}[I_0\otimes(\mathcal{U}_{31}F_1\mathcal{U}_{31}^{-1})]\}\\ &= (p')^{-1}\,\mathrm{Tr}_3\{[(I_0\otimes\mathcal{U}_{31})\rho_{01}(I_0\otimes\mathcal{U}_{31})^{-1}][(I_0\otimes\mathcal{U}_{31})(I_0\otimes F_1)(I_0\otimes\mathcal{U}_{31})^{-1}]\}\\ &= (p')^{-1}\,\mathrm{Tr}_3[(I_0\otimes\mathcal{U}_{31})\rho_{01}(I_0\otimes F_1)(I_0\otimes\mathcal{U}_{31})^{-1}]=(p')^{-1}\,\mathrm{Tr}_1[\rho_{01}(I_0\otimes F_1)].\end{aligned}$$

Using (65) again, one obtains

$$\begin{aligned}p' &\equiv \mathrm{Tr}_{03}\{\rho'_{03}[I_0\otimes(\mathcal{U}_{31}F_1\mathcal{U}_{31}^{-1})]\}\\ &= \mathrm{Tr}_{03}\{(I_0\otimes\mathcal{U}_{31})\rho_{01}(I_0\otimes\mathcal{U}_{31}^{-1})[(I_0\otimes\mathcal{U}_{31})(I_0\otimes F_1)(I_0\otimes\mathcal{U}_{31})^{-1}]\}\\ &= \mathrm{Tr}_{03}\{(I_0\otimes\mathcal{U}_{31})\rho_{01}(I_0\otimes F_1)(I_0\otimes\mathcal{U}_{31})^{-1}]\}=\mathrm{Tr}_{01}[\rho_{01}(I_0\otimes F_1)]=p.\end{aligned}$$

This establishes the validity of (70a). The proof of (70b) is analogous using (68) instead of (65). □

Corollary 25. If an observable A_{12} is suitable for deterministic teleportation, i.e. if it is complete and all its characteristic vectors are atomic state vectors, then all the results of this section are valid for deterministic teleportation via measurement of A_{12}.

Proof. Follows from the fact that, whatever the result in the measurement of A_{12}, statistical teleportation takes place. □

10. Concluding results and discussion

10.1. Why linearity or, equivalently, unitarity?

Assuming linearity of $\mathcal{U}_{31}$ (distant preparation), as we did in our derivation of the central theorem, and taking into account the fact that $\mathcal{U}_{31}$ preserves the norm, one concludes that it is unitary. Besides, since $\mathcal{U}_3 = \mathcal{I}_{31} \circ \mathcal{U}_{31}^{-1}$ (cf figure 2), and since $\mathcal{I}_{31}$ is a unitary isomorphism, also $\mathcal{U}_3$ is unitary.

One wonders if this is necessary. It might be conceivable that $\mathcal{U}_{31}$ and $\mathcal{U}_3$ violate linearity, but in way that cancels out in consecutive application (like e.g. in case of antiunitary operators). We show now that under a few assumptions that we consider essential for statistical teleportation, $\mathcal{U}_3$ and $\mathcal{U}_{31}$ *must be unitary*.

Our *assumptions* are as follows.

(i) There exists a physical system, we call it (the distant) 'laboratory', and denote it by 'L', such that the composite system $(3 + L)$ is dynamically closed during teleportation. Hence, the evolution of the system $(3+L)$ during the second step of teleportation is governed by a unitary operator $\mathcal{U}_{3L}$.

(ii) The state of the laboratory at the moment when the first step, i.e. distant preparation, is completed, is described by one and the same state vector $|\Psi\rangle'_L$ (or state operator ρ'_L) independently of the distantly prepared state vector $|\psi\rangle'_3$ $(\equiv \mathcal{U}_{31}|\psi\rangle_1, |\psi\rangle_1 \in \boldsymbol{S}_1)$.

(iii) Besides the initial state vector $|\psi\rangle'_3|\Psi\rangle'_L$ for the second step of teleportation also the final state vector $|\psi\rangle_3|\Psi\rangle_L$ has to be uncorrelated for each $|\psi\rangle'_3$ $(\in \boldsymbol{S}'_3 \equiv \mathcal{U}_{31}\boldsymbol{S}_1)$.

(iv) Both $\mathcal{U}_{31}$ and $\mathcal{U}_3$ have to be bijective (because so is the unitary isomorphism $\mathcal{I}_{31} = \mathcal{U}_3 \circ \mathcal{U}_{31}$).

(v) If $|\psi\rangle'_3, |\bar{\psi}\rangle'_3 \in \boldsymbol{S}'_3$, and $\langle\psi|'_3|\bar{\psi}\rangle'_3 = 0$, then the teleported state vectors cannot be collinear, i.e. $|\psi\rangle_3 \neq \mathrm{e}^{\mathrm{i}\lambda}|\bar{\psi}\rangle_3, \lambda \in \boldsymbol{R}$.

Violation of (v) would also violate the requirement that teleportation take distinct pure states into distinct pure states, or else, distant preparation would have to take two state vectors $|\psi\rangle_1, |\bar{\psi}\rangle_1 \in \boldsymbol{S}_1$ representing the same pure state $(|\bar{\psi}\rangle_1 = \mathrm{e}^{\mathrm{i}\omega}|\psi\rangle_1, \omega \in \boldsymbol{R})$ into orthogonal state vectors $|\psi\rangle'_3, |\bar{\psi}\rangle'_3 \in \boldsymbol{S}_3$. This is not possible, because it is obvious from the distant-preparation formula

$$|\psi\rangle'_3 \equiv (\bar{w})^{-1/2}\langle\bar{a}|_{12}|\psi\rangle_1|\bar{\psi}\rangle_{23}$$

(cf (6*a*)) that one has necessarily $|\bar{\psi}\rangle'_3 = \mathrm{e}^{\mathrm{i}\omega}|\psi\rangle'_3, \omega \in \boldsymbol{R}$, in our case.

The unitarity of $\mathcal{U}_{3L}$, in conjunction with assumptions (ii)–(iv), implies that one can fix a state vector $|\Psi\rangle_L$ of the 'laboratory' that is independent of $|\psi\rangle'_3 (\in \boldsymbol{S}'_3)$, and define an operator $\mathcal{U}_3$ mapping $\boldsymbol{S}'_3$ onto $\boldsymbol{S}_3$ that is almost-linear. This is proved in appendix B.

Further, on account of the almost-linearity of the operator $\mathcal{U}_3$, so is $\mathcal{U}_3^{-1}$, and $\mathcal{U}_{31}$ $(= \mathcal{U}_3^{-1} \circ \mathcal{I}_{31})$. This almost-linearity of $\mathcal{U}_{31}$ is proved in appendix C.

The operator $\mathcal{U}_{31}$ turns out to be 'linear with respect to' a fixed set of vectors. This is proved in appendix D. This is sufficient for it to be linear. This is proved in appendix E.

Since $\mathcal{U}_{31}$ takes state vectors into state vectors and it is linear, it is unitary; and so is $\mathcal{U}_3$ $(= \mathcal{I}_{31} \circ \mathcal{U}_{31}^{-1})$.

Finally, it is hard to imagine a macroscopic system like the one we call the (distant) 'laboratory' in a pure state because practically it is almost impossible to bring a system of great complexity into such a state. Hence, it is desirable to assume that the laboratory is in a mixed state. We generalize our proof to the case when the state vectors $|\Psi\rangle'_L$ and $|\Psi\rangle_L$ are replaced by state operators ρ'_L and ρ_L in appendix F. (In the proof we restrain from the dubious physical assumption that the individual laboratory must be in some unknown pure state.)

10.2. Is there any intermediary case between statistical and deterministic teleportation?

One may have an observable B_{12} that has some nondegenerate characteristic values with atomic state vectors as the corresponding characteristic vectors (teleporting results), but that also has degenerate or nondegenerate not teleporting characteristic values. Then, after measuring this observable, if one obtains a teleporting result, one 'phones' the distant laboratory and one informs it which result it is. There one applies the corresponding operation and the (statistical) teleportation is achieved (in a subensemble).

This is, actually, a more intricate form of statistical teleportation.

10.3. How complete is our discussion on statistical and deterministic teleportation?

To end our investigation in a self-critical way, let us sum up what has been done emphasizing what has to be done in further research.

Statistical teleportation has been treated in this work in a rather complete fashion. The general form of an atomic state vector $|\bar{a}\rangle_{12}$ has been found via (33) and (34), and the set of *all* solutions in relevant atom state vectors $|a\rangle_{12}$ has been classified in the following three ways.

(i) Classification in two steps. In the first step we consider all atom state vectors $|a\rangle_{12}$ determining one and the same intermediary space $\boldsymbol{S}'_3$ (cf (6*a*,*b*)), which then, in turn, determines $\boldsymbol{S}_2$ (cf definition 9), and, in the second step through arbitrary ON bases $\{|i\rangle_1 : i = 1, 2, \ldots, N\}$ in $\boldsymbol{S}_1$ in the Schmidt biorthogonal expansion of $|a\rangle_{12}$:

$$|a\rangle_{12} = \sum_i w^{1/2} r_i^{-1/2} |i\rangle_1 |i_2\rangle$$

where the ON basis $\{|i\rangle_2 : i = 1, 2, \ldots, N\}$ spanning $\boldsymbol{S}_2$ is an arbitrary fixed characteristic subbasis of ρ_2^B (spanning its range).

(ii) The first step is the same as in (i), and in the second classification takes place through arbitrary inverse correlation operators $\mathcal{U}_a^{-1}$ (cf (15*b*)):

$$|a\rangle_{12} = \sum_{i=1}^{N} w^{1/2} r_i^{-1/2} [(\mathcal{U}_a)^{-1} |i\rangle_2]_1 \otimes |i\rangle_2.$$

(iii) After the first step as in (i), we have classification through arbitrary unitary operators $\mathcal{U}_1$ and $\boldsymbol{S}_1$:

$$|a\rangle_{12} = (\mathcal{U}_1 \otimes I_2) |a^{(0)}\rangle_{12}$$

where $|a^{(0)}\rangle_{12}$ is a fixed atom state vector (cf corollary 15).

What we have not done in this study is a clarification which spaces $\boldsymbol{S}'_3$ $(\subseteq R[(\bar{\rho}_3^B)]^{1/2})$ do correspond to some atom state vectors $|a\rangle_{12}$ and which do not. Then, the 'good' $\boldsymbol{S}'_3$

serve in a precise way as the first step of classification, and (taking $\boldsymbol{S}_3'$ as the third input entity) the above classifications are complete.

For deterministic teleportation we have obtained necessary and sufficient conditions, but without a general solution. It is not clear whether, for an arbitrary given entangled bridge state vector $|\bar{\Psi}\rangle_{23}$, and for a given finite-dimensional subspace $\boldsymbol{S}_1$ ($\subseteq \mathcal{H}_1$) there does necessarily exist a deterministically teleporting observable A_{12}. If there exists one, it is not clear how one can classify *all* such observables A_{12}.

Acknowledgments

We express our hearty thanks to Professor M Vujičić for his kind interest and support, as well as for his suggestions for improvement of this study.

Appendix A

In connection with definition 11 (on suitability of a state operator), we prove now:

Lemma A.1. Let $\boldsymbol{S}_1$ ($\subseteq \mathcal{H}_1$) be a subspace, and let the projector P_1 project $\mathcal{H}_1$ onto it. The following four properties of a state operator ρ_1 are equivalent to each other.

(i) There exists a decomposition of ρ_1 into pure states, $\rho_1 = \sum_k w_k |\psi_k\rangle_1\langle\psi_k|_1$, such that $\forall k : |\psi_k\rangle_1 \in \boldsymbol{S}_1$.

(ii) $P_1\rho_1 = \rho_1$.

(iii) The subspace $\boldsymbol{S}_1$ is invariant for ρ_1 and the latter reduces into zero in its orthocomplement $\boldsymbol{S}_1^{\perp}$.

(iv) For every decomposition of ρ_1 into pure states, each state vector obtained in the decomposition (cf (i)) belongs to $\boldsymbol{S}_1$.

Proof. The relation $|\psi_k\rangle_1 \in \boldsymbol{S}_1$ is equivalent to $P_1|\psi_k\rangle_1 = |\psi_k\rangle_1$. Hence, (i) obviously implies (ii). Adjoining the latter, we see that P_1 and ρ_1 commute. Hence, $\boldsymbol{S}_1$ and $\boldsymbol{S}_1^{\perp}$ are invariant for ρ_1. If $|\phi\rangle_1 \in \boldsymbol{S}_1^{\perp}$, then, also $(\rho_1|\phi\rangle_1) \in \boldsymbol{S}_1^{\perp}$. On the other hand, as a consequence of (ii), $P_1(\rho_1|\phi\rangle_1) = (\rho_1|\phi\rangle_1)$. Since $\boldsymbol{S}_1^{\perp} = (R(P_1))^{\perp}$, the last two relations are possible only if $(\rho_1|\phi\rangle_1) = 0$. Thus, (ii) implies (iii).

We assume that (iii) is valid, and we take an orthonormal subbasis $\{|\phi_j\rangle_1 : \forall j\}$ spanning $\boldsymbol{S}_1^{\perp}$. Let $\rho_1 = \sum_k w_k|\psi_k\rangle_1\langle\psi_k|_1$ be an arbitrary decomposition of ρ_1 into pure states. Then,

$$\forall j : 0 = \langle\phi_j|\rho_1|\phi_j\rangle_1 = \sum_k w_k \langle\phi_j|_1|\psi_k\rangle_1\langle\psi_k|_1|\phi_j\rangle_1.$$

Hence, $\forall j, \forall k : \langle\phi_j|\psi_k\rangle_1 = 0$, and this implies that $\forall k : |\psi_k\rangle_1 \in \boldsymbol{S}_1$. Thus, (iii) implies (iv). Finally, (i) is a consequence of (iv) because there always exists the spectral decomposition of ρ_1 into characteristic vectors (and w_k are then the corresponding characteristic values). □

Appendix B

Lemma B.1. One can fix a state vector $|\Psi\rangle_L$ of the 'laboratory' that is independent of $|\psi\rangle_3'$ ($\in \boldsymbol{S}_3'$), and such that

$$\forall |\psi\rangle_3' \in \boldsymbol{S}_3': \qquad \mathcal{U}_{3L}(|\psi\rangle_3' \otimes |\Psi\rangle_L') \equiv (\mathcal{U}_3|\psi\rangle_3') \otimes |\psi\rangle_L$$

where $\mathcal{U}_3$ is almost-linear in the sense that for a certain choice of an ON basis $\{|k\rangle_3' : \forall k\}$ spanning $\boldsymbol{S}_3'$ one has

$$|\psi\rangle_3' = \sum_k \alpha_k |k\rangle_3' \Rightarrow \mathcal{U}_3|\psi\rangle_3' = \sum_k \alpha_k \mathrm{e}^{\mathrm{i}\lambda_k}(\mathcal{U}_3|k\rangle_3') \tag{B.1}$$

($\lambda_k \in \mathbf{R}$, and they depend on the corresponding $|k\rangle_3'$, but not on $|\psi\rangle_3'$).

Before we prove lemma B.1, we insert an auxiliary result.

Lemma B.2. If $|\psi\rangle_3', |\bar{\psi}\rangle_3' \in \boldsymbol{S}_3'$, and $|\psi\rangle_3', |\bar{\psi}\rangle_3'$ are state vectors orthogonal to each other, then $|\Psi\rangle_L = \mathrm{e}^{\mathrm{i}\lambda}|\bar{\Psi}\rangle_L, \lambda \in \mathbf{R}$, where $\mathcal{U}_{3L}(|\psi\rangle_3'|\Psi\rangle_3') = |\psi\rangle_3|\Psi\rangle_L$ and $\mathcal{U}_{3L}(|\bar{\psi}\rangle_3|\Psi\rangle_3') = |\bar{\psi}\rangle_3|\bar{\Psi}\rangle_L$.

Proof. Since the state vectors $|\psi\rangle_3', |\bar{\psi}\rangle_3'$ are orthogonal, so are $|\psi\rangle_3'|\Psi\rangle_L'$ and $|\bar{\psi}\rangle_3'|\Psi\rangle_L'$. Then, on account of the unitarity of $\mathcal{U}_{3L}$ (cf assumption (i)), so are also $|\psi\rangle_3|\Psi\rangle_L (= \mathcal{U}_{3L}(|\psi\rangle_3'|\Psi\rangle_L'))$ and $|\bar{\psi}\rangle_3|\bar{\Psi}\rangle_L (= \mathcal{U}_{3L}(|\bar{\psi}\rangle_3'|\Psi\rangle_L'))$. Hence, either $\langle\Psi|_L|\bar{\Psi}\rangle_L = 0$ or $\langle\psi|_3|\bar{\psi}\rangle_3 = 0$.

(a) Let $\langle\Psi|_L|\bar{\Psi}\rangle_L = 0$. Then, due to assumptions (i) and (ii), one has:

$$\mathcal{U}_{3L}(|\psi\rangle_3'|\Psi\rangle_L' + |\bar{\psi}\rangle_3'|\Psi\rangle_L') = |\psi\rangle_3|\Psi\rangle_L + |\bar{\psi}\rangle_3|\bar{\Psi}\rangle_L. \tag{B.2}$$

On the other hand, LHS $= \mathcal{U}_{3L}[(|\psi\rangle_3' + |\bar{\psi}\rangle_3') \otimes |\Psi\rangle_L']$, and owing to assumption (iii)

$$\text{LHS} = p|x\rangle_3|\Psi_x\rangle_L \tag{B.3}$$

where p is the norm of the LHS. Relation (B.2) then gives $p^2 = 2$. Further, (B.2) and (B.3) imply

$$|\Psi_x\rangle_L = 2^{-1/2}(\langle x|_3|\psi\rangle_3)|\Psi\rangle_L + (\langle x|_3|\bar{\psi}\rangle_3)|\bar{\Psi}\rangle_L. \tag{B.4}$$

Since in (B.4) all kets are vectors of norm 1 and by assumption (a) $\langle\Psi|_L|\bar{\Psi}\rangle_L = 0$, necessarily

$$|\langle x|_3|\psi\rangle_3| = |\langle x|_3|\bar{\psi}\rangle_3| = 1$$

or

$$\mathrm{e}^{\mathrm{i}\mu}|\psi\rangle_3 = |x\rangle_3 = \mathrm{e}^{\mathrm{i}\omega}|\bar{\psi}\rangle_3 \qquad \mu, \omega \in \mathbf{R}.$$

Owing to the assumption $\langle\psi|_3'|\bar{\psi}\rangle_3' = 0$, and to assumption (v) in the text, this is not possible.

(b) Let $\langle\psi|_3|\bar{\psi}\rangle_3 = 0$. Since the two tensor factors play symmetrical roles, one proves in analogy with (a) that

$$\mathrm{e}^{\mathrm{i}\mu}|\Psi\rangle_L = |\Psi_x\rangle_L = \mathrm{e}^{\mathrm{i}\omega}|\bar{\Psi}\rangle_L \qquad \mu, \omega \in \mathbf{R}.$$

□

Proof. Proof of lemma B.1. We assume that we have two arbitrary state vectors $|\psi\rangle_3', |1\rangle_3' \in \boldsymbol{S}_3'$, and we complete the second one into an orthonormal basis $\{|k\rangle_3' : k = 1, 2, \ldots, N\}$ spanning $\boldsymbol{S}_3'$. Finally, let $|\psi\rangle_3' = \sum_k \alpha_k |k\rangle_3'$. Then

$$\mathcal{U}_{3L}(|\psi\rangle_3'|\Psi\rangle_L') = |\psi\rangle_3|\Psi\rangle_L$$

and, on the other hand, since $\mathcal{U}_{3L}$ is linear, utilizing lemma B.2, one can write

$$\text{LHS} = \sum_k \alpha_k \mathcal{U}_{3L}(|k\rangle_3'|\Psi\rangle_L') = \sum_k \alpha_k |k\rangle_3|\Psi_k\rangle_L = (\alpha_1|1\rangle_3 + \sum_{k=2}^{N} \alpha_k \mathrm{e}^{\mathrm{i}\lambda_k}|k\rangle_3)|\Psi_1\rangle_L.$$

Hence, we can define

$$|\Psi\rangle_L \equiv |\Psi_1\rangle_L \qquad |\psi\rangle_3 \equiv \mathcal{U}_3|\psi\rangle_3' \equiv \sum_{k=1}^{N} \alpha_k \mathrm{e}^{\mathrm{i}\lambda_k}|k\rangle_3$$

with

$$\forall k: \qquad \mathcal{U}_{3L}(|k\rangle_3'|\Psi\rangle_L') = |k\rangle_3|\Psi\rangle_L \equiv (\mathcal{U}_3|k\rangle_3')|\Psi\rangle_L$$

and $\lambda_1 \equiv 0$. □

Appendix C

The inverse $\mathcal{U}_3^{-1}$ exists due to assumption (iv) (cf section 10.1). First, we prove that it is almost-linear with respect to the set of vectors $\{|k\rangle_3 \equiv \mathcal{U}_3|k\rangle_3' : \forall k\}$ in the span of these vectors. Utilizing (B.1), one obtains:

$$|\psi\rangle_3 \equiv \sum_k \beta_k|k\rangle_3 = \sum_k \beta_k(\mathcal{U}_3|k\rangle_3') = \mathcal{U}_3\Big(\sum_k \beta_k \mathrm{e}^{-\mathrm{i}\lambda_k}|k\rangle_3'\Big)$$

hence,

$$\mathcal{U}_3^{-1}|\psi\rangle_3 = \sum_k \beta_k \mathrm{e}^{-\mathrm{i}\lambda_k}(\mathcal{U}_3^{-1}|k\rangle_3). \tag{C.1}$$

Next, we prove that $\mathcal{U}_{31}$ $(= \mathcal{U}_3^{-1} \circ \mathcal{I}_{31})$ is almost-linear with respect to the set of vectors $\{|k\rangle_1 \equiv \mathcal{U}_{31}^{-1}|k\rangle_3' : \forall k\}$ in their span. The inverse $\mathcal{U}_{31}^{-1}$ exists on account of assumption (iv). Let $|\psi\rangle_1 \equiv \sum_k \beta_k|k\rangle_1$. Then, owing to (C.1),

$$\mathcal{U}_{31}|\psi\rangle_1 = (\mathcal{U}_3^{-1} \circ \mathcal{I}_{31})|\psi\rangle_1 = \mathcal{U}_3^{-1}\Big(\sum_k \beta_k|k\rangle_3\Big) = \sum_k \beta_k \mathrm{e}^{-\mathrm{i}\lambda_k}|k\rangle_3' = \sum_k \beta_k \mathrm{e}^{-\mathrm{i}\lambda_k}(\mathcal{U}_{31}|k\rangle_1). \tag{C.2}$$

□

Appendix D

By a variation of the necessity-part of the proof of lemma 8, we first prove that $\mathcal{U}_{31}$ satisfying (C.2) is linear with respect to the set of vectors $\{|k\rangle_1 \equiv \mathcal{U}_{31}^{-1}|k\rangle_3' : \forall k\}$ i.e. that

$$|\psi\rangle_1 \equiv \sum_k \beta_k|k\rangle_1 \Rightarrow \mathcal{U}_{31}|\psi\rangle_1 = \sum_k \beta_k(\mathcal{U}_{31}|k\rangle_1). \tag{D.1}$$

Let the probabilities of distant preparation be w and $\{w_k : \forall k\}$ respectively, i.e.

$$\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}) = w^{1/2}|\psi\rangle_3' \tag{D.2a}$$

$$\forall k: \qquad \langle a|_{12}(|k\rangle_1|\Psi\rangle_{23}) = w_k^{1/2}|k\rangle_3'. \tag{D.2b}$$

Relation (D.2a) can be rewritten in the form

$$\langle a|_{12}(|\psi\rangle_1|\Psi\rangle_{23}) \equiv w^{1/2}(\mathcal{U}_{31}|\psi\rangle_1).$$

Replacing here $|\psi\rangle_1$ in its expanded form (cf (D.1)), and utilizing (D.2b) and (C.2), one obtains

$$\sum_k \beta_k w_k^{1/2}|k\rangle_3' = w^{1/2}\sum_k \beta_k \mathrm{e}^{-\mathrm{i}\lambda_k}|k\rangle_3'. \tag{D.3}$$

Since $\{|k\rangle_3' : \forall k\}$ was taken to be an ON basis in $\boldsymbol{S}_3'$ (cf appendix B), (D.3) implies

$$\forall k : \beta_k w_k^{1/2} = \beta_k w^{1/2} \mathrm{e}^{-\mathrm{i}\lambda_k}. \tag{D.4}$$

Choosing $|\psi\rangle_1$ so that all expansion coefficients are nonzero (cf (D.1)), (D.4) entails: $\forall k : \mathrm{e}^{-\mathrm{i}\lambda k} = 1$. Substituting this in (C.2), we derive the claim of (D.1).

Appendix E

Lemma E.1. If an operator $\mathcal{U}_{31}$, mapping $\boldsymbol{S}_1$ into $\boldsymbol{S}_3'$, is 'linear with respect to' a set of vectors $\{|k\rangle_1 : \forall k\}$, then it is linear in the span of these vectors.

Proof. Let $|\psi\rangle_1$ and $\{|\psi_j\rangle_1 : j = 1, 2, \ldots, J\}$ be in the span of the above vectors, i.e. let it be possible to expand them (at least in one way):

$$|\psi\rangle_1 = \sum_k \beta_k |k\rangle_1 \tag{E.1a}$$

$$\forall_j: \qquad |\psi_j\rangle_1 = \sum_k \gamma_{jk} |k\rangle_1. \tag{E.1b}$$

Besides, let

$$|\psi\rangle_1 = \sum_{j=1}^{J} \alpha_j |\psi_j\rangle_1. \tag{E.2}$$

Substituting (E.1) in (E.2), we further have

$$|\psi\rangle_1 = \sum_{j=1}^{J} \sum_k \alpha_j \gamma_{jk} |k\rangle_1.$$

On account of the 'linearity with respect to' the set $\{|k\rangle_1 : \forall k\}$, one has

$$\mathcal{U}_{31}|\psi\rangle_1 = \sum_{j=1}^{J} \sum_k \alpha_j \gamma_{jk} (\mathcal{U}_{31}|k\rangle_1). \tag{E.3}$$

On the other hand, applying $\mathcal{U}_{31}$ to (E.1b), one obtains

$$\mathcal{U}_{31}|\psi_j\rangle_1 = \sum_k \gamma_{jk} (\mathcal{U}_{31}|k\rangle_1).$$

Replacing this in (E.3), one arrives at

$$\mathcal{U}_{31}|\psi\rangle_1 = \sum_{j=1}^{J} \alpha_j (\mathcal{U}_{31}|\psi_j\rangle_1).$$

In view of the arbitrariness of the linear combination (E.2), this proves linearity. □

Appendix F

Lemma F.1. The final state operator ρ_L of the 'laboratory' is independent of the distantly-prepared state vector $|\psi\rangle'_3$.

Proof. The assumptions on the second step of statistical teleportation (see section 10.1) allow us to write

$$|\psi\rangle_3\langle\psi|_3 \otimes \rho_L = \mathcal{U}_{3L}(|\psi\rangle'_3\langle\psi|'_3 \otimes \rho'_L)\mathcal{U}^\dagger_{3L}$$

where ρ_L might depend on $|\psi\rangle'_3$. Let

$$\rho'_L = \sum_i w_i |\Psi_i\rangle'_L\langle\Psi_i|'_L$$

be a spectral form of ρ'_L. Substitution of the latter relation in the former gives

$$|\psi\rangle_3\langle\psi|_3 \otimes \rho_L = \sum_i w_i \mathcal{U}_{3L}(|\psi\rangle'_3\langle\psi|'_3 \otimes |\Psi_i\rangle'_L\langle\Psi_i|'_L)\mathcal{U}^\dagger_{3L}. \tag{F.1}$$

Since for any value of i, $|\Psi_i\rangle'_L$ is independent of $|\psi\rangle'_3$ (because so is ρ'_L by assumption (ii) in the text), according to lemma B.1, we have

$$\forall i : \mathcal{U}_{3L}(|\psi\rangle'_3\langle\psi|'_3 \otimes |\Psi_i\rangle'_L\langle\Psi_i|'_L)\mathcal{U}^\dagger_{3L} = (|\psi\rangle_3\langle\psi|_3 \otimes |\Psi_i\rangle_L\langle\Psi_i|_L)$$

and the final state $|\Psi_i\rangle_L\langle\Psi_i|_L$ is independent of $|\psi\rangle'_3$.

Replacement of this relation in (F.1) results in

$$|\psi\rangle_3\langle\psi|_3 \otimes \rho_L = \sum_i w_i(|\psi\rangle_3\langle\psi|_3 \otimes |\Psi_i\rangle_L\langle\Psi_i|_L) = |\psi\rangle_3\langle\psi|_3 \otimes \left(\sum_i w_i |\Psi_i\rangle_L\langle\Psi_i|_L\right).$$

□

References

Bell J S and Nauenberg M 1966 *Preludes in Theoretical Physics* (Amsterdam: North-Holland) p 279

Bennett Ch H 1995 *Phys. Today* **October** 24

Bennett Ch H, Brassard G, Crépeau C, Jozsa R, Peres A and Wootters W K 1993 *Phys. Rev. Lett.* **70** 1895

Bennett Ch H, Brassard G, Popescu S, Schumacher B, Smolin J A and Wootters W K 1996 *Phys. Rev. Lett.* **76** 722

Braunstein S L and Mann A 1995 *Phys. Rev.* **A51** R1727

Busch P, Lahti P and Mittelstaedt P 1991 *Quantum Theory of Measurement* (Berlin: Springer)

D'Espagnat B 1976 *Conceptual Foundations of Quantum Mechanics* (Reading, MA: Benjamin)

Elitzur A C and Vaidman L 1993 *Found. Phys.* **23** 987

Gudder S P 1979 *Stochastic Methods in Quantum Mechanics* (New York: North-Holland)

Herbut F 1974 *Int. J. Theor. Phys.* **11** 193

——1986 *Int. J. Theor. Phys.* **25** 1215

Herbut F and Vujičić M 1976 *Ann. Phys., NY* **96** 382

Kwiat P G, Weinfurter H, Herzog T, Zeilinger A and Kasevich M A 1995 *Phys. Rev. Lett.* **74** 4763

Kwiat P G, Weinfurter H and Zeilinger A 1996 *Sci. Am.* **November** 52

Lüders G 1951 *Ann. Phys., Lpz.* **8** 322

Messiah A 1961 *Quantum Mechanics* vol 1 (Amsterdam: North-Holland) p 333

Moussa M H Y 1997 *Phys. Rev.* A **55** R3287

Nielsen M A and Caves C M 1997 *Phys. Rev.* A **55** 2547

Pauli W 1933 *Die Allgemeinen Prinzipien der Wellenmechanik (Handbuch der Physik XXIV)* (Berlin: Springer)

Popescu S 1994 *Phys. Rev. Lett.* **72** 797

Schrödinger E 1936 *Proc. Camb. Phil. Soc.* **32** 446

Vujičić M and Herbut F 1984 *J. Math. Phys.* **25** 2253

Vujičić M and Herbut F 1988 *J. Phys. A: Math. Gen.* **21** 2931

universe

MDPI

Article

Equivalence Principle in Classical and Quantum Gravity

Nikola Paunković [1,†] and Marko Vojinović [2,*,†]

[1] Instituto de Telecomunicações and Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal
[2] Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia
* Correspondence: vmarko@ipb.ac.rs
† These authors contributed equally to this work.

Abstract: We give a general overview of various flavours of the equivalence principle in classical and quantum physics, with special emphasis on the so-called weak equivalence principle, and contrast its validity in mechanics versus field theory. We also discuss its generalisation to a theory of quantum gravity. Our analysis suggests that only the strong equivalence principle can be considered fundamental enough to be generalised to a quantum gravity context since all other flavours of equivalence principle hold only approximately already at the classical level.

Keywords: equivalence principle; general relativity; quantum gravity

check for updates

Citation: Paunković, N.; Vojinović, M. Equivalence Principle in Classical and Quantum Gravity. *Universe* **2022**, *8*, 598. https://doi.org/10.3390/universe8110598

Academic Editors: Arundhati Dasgupta and Alfredo Iorio

Received: 1 October 2022
Accepted: 10 November 2022
Published: 12 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Quantum mechanics (QM) and general relativity (GR) are the two cornerstones of modern physics. Yet, merging them together in a quantum theory of gravity (QG) is still elusive despite the nearly century-long efforts of vast numbers of physicists and mathematicians. While the majority of the attempts were focused on trying to formulate the full theory of quantised gravity, such as string theory, loop quantum gravity, non-commutative geometry, and causal set theory, to name a few, a number of recent studies embraced a rather more modest approach by exploring possible consequences of basic features and principles of QM and GR, and their status, in a tentative theory of QG. Acknowledging that the superposition principle, as a defining characteristic of any quantum theory, must be featured in QG as well, led to a number of papers studying gravity-matter entanglement [1–7], genuine indefinite causal orders [8–15], quantum reference frames [16–20] and deformations of Lorentz symmetry [21–25], to name a few major research directions. Exploring spatial superpositions of masses, and in general gravitational fields, led to the analysis of the status of various versions of the equivalence principle, and their exact formulations in the context of QG. In particular, in [26], it was shown that the weak equivalence principle (WEP) should generically be violated in the presence of a specific type of superpositions of gravitational fields, describing small quantum fluctuations around a dominant classical geometry. On the other hand, a number of recent studies propose generalisations of WEP to QG framework (see for example [16,20,27–31]), arguing that it remains satisfied in such scenarios, a result *seemingly* at odds with [26] (for details, see the discussion from Section 5).

The modern formulation of WEP is given in terms of a *test particle* and it's *trajectory*: it is a *theorem* within the mathematical formulation of GR stating that the trajectory of a test particle satisfies the so-called geodesic equation [32–46], while it is *violated* within the context of QG, as shown in [26]. In this paper, we present a brief overview of WEP in GR and a critical analysis of the notions of particle and trajectory in both classical and quantum mechanics, as well as in the corresponding field theories.Our analysis demonstrates that WEP, as well as all other flavours of the equivalence principle (EP) aside from the strong one (SEP), hold only approximately. From this we conclude that neither WEP nor any other flavour of EP (aside from SEP) can be considered a viable candidate for generalisation to the quantum gravity framework.

The paper is organised as follows. In Section 2, we give a brief historical overview of various flavours of the equivalence principle, with a focus on WEP. In Section 3, we analyse the notion of a trajectory in classical and quantum mechanics, while in Section 4 we discuss the notion of a particle in field theory and QG. Finally, in the Conclusion, we briefly review and discuss our results, and present possible future lines of research.

2. Equivalence Principle in General Relativity

The equivalence principle is one of the most fundamental principles in modern physics. It is one of the two cornerstone building blocks for GR, the other being the principle of general relativity. While its importance is well understood in the context of gravity, it is often underappreciated in the context of other fundamental interactions. In addition, there have been numerous studies and everlasting debates about whether EP holds also in quantum physics, if it should be generalised to include quantum phenomena or not, etc. Finally, EP has been historically formulated in a vast number of different ways, which are often not mutually equivalent, leading to a lot of confusion about the actual statement of the principle and its physical content [47–53]. Given the importance of EP, and the amount of confusion around it, it is important to try and help clarify these issues.

The equivalence principle is best introduced by stating its purpose—in its traditional sense, the purpose of EP is to *prescribe the interaction between gravity and all other fields in nature, collectively called matter* (by "matter" we assume not just fermionic and scalar fields, but also gauge vector bosons, i.e., nongravitational interaction fields). This is important to state explicitly since EP is often mistakenly portrayed as a property of gravity alone, without any reference to matter. In a more general, less traditional, and often not appreciated sense, the purpose of EP is to prescribe the interaction between *any gauge field* and all other fields in nature (namely fermionic and scalar matter, as well as other gauge fields, including gravity), which we will reflect on briefly in the case of electrodynamics below.

Given such a purpose, let us for the moment concentrate on the gravitational version of EP, and provide its modern formulation, as it is known and understood today. The statement of the equivalence principle is the following:

> *The equations of motion for matter coupled to gravity remain locally identical to the equations of motion for matter in the absence of gravity.*

This kind of statement requires some unpacking and comments.

- When comparing the equations of motion in the presence and in the absence of gravity, the claim that they remain identical may naively suggest that gravity does not influence the motion of matter in any way whatsoever. However, on closer inspection, the statement is that the two sets of equations remain *locally* identical, emphasising that the notion of locality is a crucial feature of the EP. While equations of motion are already local in nature (since they are usually expressed as partial differential equations of finite order), the actual interaction between matter and gravity enters only when *integrating* those equations to find a solution (see Appendix A for a detailed example).
- In order to compare the equations of motion for matter in the presence of gravity to those in its absence, the equations themselves need to be put in a suitable form (typically expressed in general curvilinear coordinates, as tensor equations). The statement of EP relies on a theorem that this can always be achieved, first noted by Erich Kretschmann [54].
- Despite being dominantly a statement about the interaction between matter and gravity, EP also implicitly suggests that the best way to describe the gravitational field is as a property of the geometry of spacetime, such as its metric [55]. In that setup, EP can be reformulated as a statement of *minimal coupling* between gravity and matter, stating that equations of motion for matter may depend on the spacetime metric and its first derivatives, but not on its (antisymmetrised) second derivatives, i.e., the *spacetime curvature does not explicitly appear in the equations of motion for matter*.

- The generalisation of EP to other gauge fields is completely straightforward, by replacing the role of gravity with some other gauge field, and suitably redefining what matter is. For example, in electrodynamics, the EP can be formulated as follows:

  > *The equations of motion for matter coupled to the electromagnetic field remain locally identical to the equations of motion for matter in the absence of the electromagnetic field.*

  This statement can also be suitably reformulated as the minimal coupling between the electromagnetic (EM) field and matter, i.e., coupling matter to the electromagnetic potential A_μ but not to the corresponding field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This is in fact the standard way the EM field is coupled to matter (see Appendix A for an illustrative example). Even more generally, the gauge field sector of the whole Standard Model of elementary particles (SM) is built using the minimal coupling prescription, meaning that the suitably generalised version of the EP actually prescribes the interaction between matter and all fundamental interactions in nature, namely strong, weak, electromagnetic and gravitational. In this sense, EP is a cornerstone principle for the whole fundamental physics, as we understand it today.

Of course, much more can be said about the statement of EP, its consequences, and various other details. However, in this work, our attention will focus on the so-called *weak equivalence principle* (WEP), which is a reformulation of EP applied to matter which consists of mechanical particles. To that end, it is important to understand various flavours and reformulations of EP that have appeared through history.

As with any deep concept in physics, EP has been expressed historically through a painstaking cycle of formulating it in a precise way, understanding the formulation, understanding the drawbacks of that formulation, coming up with a better formulation, and repeating. In this sense, EP, as quoted above, is a modern product of long and meticulous refinement over several generations of scientists. Needless to say, each step in that process made its way into contemporary physics textbooks, leading to a plethora of different formulations of EP that have accumulated in the literature over the years. This can bring about a lot of confusion about what EP actually states [47–50] since various formulations from old and new literature may often be not merely phrased differently, but in fact substantively inequivalent. To that end, let us comment on several of the most common historical statements of EP (for a more detailed historical overview and classification, see [56,57]), and their relationship with the modern version:

- *Equality of gravitational and inertial mass.* This is one of the oldest variants of EP, going back to Newton's law of universal gravitation. The statement claims that the "gravitational charge" of a body is the same as the body's resistance to acceleration, in the sense that the mass appearing on the left-hand side of Newton's second law of motion exactly cancels the mass appearing in Newton's gravitational force law on the right-hand side. This allows one to relate it to the modern version of EP, in the sense that a suitably accelerated observer could rewrite Newton's law of motion as the equation for a free particle, exploiting the cancellation of the "intertial force" and the gravitational force on the right-hand side of the equation. Unfortunately, this version of EP is intrinsically nonrelativistic, and applicable only in the context of Newtonian gravity since already in GR the source of gravity becomes the full stress-energy tensor of matter fields, rather than just the total mass. Finally, this principle obviously fails when applied to photons, as demonstrated by the gravitational bending of light.
- *Universality of free fall.* Going back all the way to Galileo, this statement claims that the interaction between matter and gravity does not depend on any intrinsic property of matter itself, such as its mass, angular momentum, chemical composition, temperature, or any other property, leading to the idea that gravity couples universally (i.e., in the same way) to all matter. Formulated from experimental observations by Galileo, its validity is related to the quality of experiments used to verify it. As we shall see below,

in a precise enough setting, one can experimentally observe direct coupling between the angular momentum of a body and spacetime curvature [32–46], invalidating the statement.

- *Local equality between gravity and inertia.* Often called Einstein's equivalence principle, the statement claims that a local and suitably isolated observer cannot distinguish between accelerating and being at rest in a uniform gravitational field. While this statement is much closer in spirit to the modern formulation of EP, it obscures the crucial aspect of the principle — coupling of matter to gravity. Namely, in this formulation, it is merely implicit that the only way an observer can *attempt to distinguish* gravity from inertia is by making local experiments using some form of *matter*, i.e., studying the equations of motion of matter in the two situations and trying to distinguish them by observing whether or not matter behaves differently. Moreover, the statement is often discussed in the context of mechanics, arguing that any given particle does not distinguish between gravity and inertia. This has two main pitfalls—first, the reliance on particles is very misleading (as we will discuss below in much more detail), and second, it implicitly suggests that gravity and inertia are the same phenomenon, which is completely false. Namely, inertia can be understood as a specific form of gravity, but a general gravitational field cannot be simulated by inertia, since inertia cannot account for tidal effects of inhomogeneous configurations of gravity.
- *Weak equivalence principle.* Stating that the equations of motion of particles do not depend on spacetime curvature, or equivalently, that the motion of a free particle is always a geodesic trajectory in spacetime, WEP is in fact an application of modern EP to mechanical point-like particles (i.e., test particles). One can argue that, as far as the notion of a point-like particle is a well-defined concept in physics, WEP is a good principle. Nevertheless, as we will discuss below in detail, the notion of a point-like particle is an idealisation that does not actually have any counterpart in reality, in either classical or quantum physics. Regarding a realistic particle (with nonzero size), WEP *never holds*, due to the explicit effect of gravitational tidal forces across the particle's size. In this sense, WEP can be considered at best an *approximate* principle, which can be assumed to hold only in situations where particle size can be approximated to zero.
- *Strong equivalence principle.* This version of the principle states that the equations of motion of all fundamental fields in nature do not depend on spacetime curvature (see [55], Section 16.2, page 387). To the best of our knowledge so far, fields are indeed the most fundamental building blocks in modern physics (such as SM), while the strength of the gravitational field is indeed described by spacetime curvature (as in GR). In this sense, the statement of SEP is actually an instance of EP applied to field theory, and as such equivalent to the modern statement of EP. So far, all our knowledge of natural phenomena is consistent with the validity of SEP.

As can be seen from the above review, various formulations of EP are both mutually inequivalent and have different domains of applicability. Specifically, only SEP holds universally, while all other flavours of EP hold only approximately. In the remainder of the paper, we focus on the study of WEP since recently it gained a lot of attention in the literature [20,27–29,31], primarily in the context of its generalisation to a "quantum WEP", and in the context of a related question of particle motion in a quantum superposition of different gravitational configurations, the latter being a scenario that naturally arises in QG. Since WEP is stated in terms of a test particle and its trajectory, in order to try and generalise it to the scope of QG one should first analyse these two notions in classical and quantum mechanics and field theory in more detail.

3. The Notion of Trajectory in Classical and Quantum Mechanics

A trajectory of a physical system in three-dimensional space is a set of points that form a line, usually parameterised by time. More formally, a trajectory is a set $\{(x(t), y(t), z(t)) \in$

$\mathbb{R}^3 | t \in [t_i, t_f] \subset \mathbb{R} \wedge t_i < t_f\}$, given by three smooth functions $x, y, z : \mathbb{R} \mapsto \mathbb{R}$. Depending on the nature of the system, the choice of points that form its trajectory may vary.

In classical mechanics, one often considers an ideal "point-like particle" localised in one spatial point $(x(t), y(t), z(t))$ at each moment of time t, in which case the choice of the points forming the system's trajectory is obvious. In the case of systems continuously spread over certain volumes ("rigid bodies", or "objects") or composite systems consisting of several point-like particles or bodies, it is natural to consider their centres of mass as points that form the trajectory. While this definition is natural, widely accepted, and formally applicable to any classical mechanical system, there are cases in which the very notion of a trajectory loses its intuitive, as well as useful, meaning.

Consider for example an electrical dipole, i.e., a system of two point-like particles with equal masses and opposite electrical charges, separated by the distance $\ell(t)$. As long as this distance stays "small" and does not vary significantly with time, the notion of a trajectory of a dipole, defined as the set of centres of mass of the two particles, does meet our intuition, and can be useful. Informally, if the trajectories of each of the two particles are "close" to each other, they can be approximated, and consequently represented, by the trajectory of the system's centre of mass. However, if the separate trajectories of the two particles diverge, one going to the "left", and another to the "right", one could hardy talk of a trajectory of such a composite system, although the set of locations of its centres of mass is still well defined. In fact, the dipole itself ceases to make physical sense when the distance between its constituents is large.

Moving to the realm of quantum mechanics, due to the superposition principle, even the ideal point-like particles do not have a well-defined position, which is further quantified by the famous Heisenberg uncertainty relations. Thus, the trajectory of point-like particles (and any system that in a given regime can be approximated to be point-like) is defined as a set of expectation values of the position operator. Like in the case of composite classical systems, here as well the definition of a trajectory of a point-like particle is mathematically always well defined, yet for a very similar reason is applicable only to certain cases. Namely, in order to give a useful meaning to the above definition of trajectory, the system considered must be *well localised*. Consider for example the double-slit experiment, in which the point-like particle is highly delocalised so that we say that *its trajectory is not well defined*, even though the set of the expectation values of the position operator is.

We see that, while in mechanics both the notions of a particle and its trajectory are rather straightforward and always well defined, the latter make sense only if our system is well localised in space (see for example [58], where the authors analyse the effects of wave-packet spreading to the notion of a trajectory).

4. The Notion of a Particle in Field Theory

While in classical mechanics a point-like particle is always well localised, we have seen that in the quantum case one must introduce an additional constraint in order for it to be considered localised—the particle should be represented by a wave-packet. The source for this requirement lies in the fact that quantum particles, although mechanical, are represented by a *wavefunction*. Thus, it is only to be expected that when moving to the realm of the field ontology, the notion of a particle becomes even more involved and technical.

In field theory, the fundamental concept is the *field*, rather than a particle. The notion of a particle is considered a derived concept, and in fact in QFT one can distinguish two vastly different phenomena that are called "particles".

The first notion of a particle is an elementary excitation of a free field. For example, the state

$$|\Psi\rangle = \hat{a}^\dagger(\vec{k})|0\rangle\,,$$

is called a *single particle state* of the field, or a *plane-wave-particle*. It has the following properties:

- It is an eigenstate of the *particle number operator* for the eigenvalue 1.

- It has a sharp value of the momentum $\vec{k}$, and corresponds to a completely delocalised plane wave configuration of the field.
- It has no centre of mass, and no concept of "position" in space since the "position operator" is not a well-defined concept for the field.
- States of this kind are said to describe *elementary particles*, understood as asymptotic free states of past and future infinity, in the context of the S-matrix for scattering processes. An example of a real scalar particle of this type would be the *Higgs particle*. For fields of other types (Dirac fields, vector fields, etc.) examples would be an *electron*, a *photon*, a *neutrino*, an asymptotically free *quark*, and so on. Essentially, all particles tabulated in the Standard Model of elementary particles are of this type.

Note that all the above notions are defined within the scope of free field theory, and do not carry over to interacting field theory. In other words, free field theory is a convenient idealisation, which does not really reflect realistic physics. One should therefore understand the concept of a plane-wave-particle in this sense, merely as a convenient mathematical approximation. Moreover, the particle number operator is not an invariant quantity, as demonstrated by the Unruh effect. We should also emphasise that in an interacting QFT, the proper way to understand the notion of a particle is as a localised wave-packet, interacting with its virtual particle cloud, which does have a position in space and whose momentum is defined through its group velocity. In this sense, the particle as a wave-packet could be better interpreted as a kink, discussed below.

The second notion of the particle in field theory is a bound state of fields, also called a *kink solution*. This requires an interacting theory since interactions are necessary to form bound states. This kind of configuration of fields has the following properties:

- It is not an eigenstate of the particle number operator, and the expectation value of this operator is typically different from 1.
- It is usually well localised in space, and does not have a sharp value of momentum.
- As long as the kink maintains a stable configuration (i.e., as long as it does not decay), one can in principle assign to it the concept of *size*, and as a consequence also the concepts of *centre of mass*, *position in space*, and *trajectory*. In this sense, a kink can play the role of a test particle.
- States of this kind are said to describe *composite particles*. Given an interacting theory such as the Standard Model, under certain circumstances quarks and gluons form bound states called a *proton* and a *neutron*. Moreover, protons and neutrons further form bound states called *atomic nuclei*, which together with electrons and photons form *atoms*, *molecules*, and so on.

For a kink, the notions of centre of mass, position in space and size are described only as classical concepts, i.e., as expectation values of certain field operators, such as the stress-energy tensor. Moreover, given the nonzero size of the kink, its centre of mass and position are not uniquely defined, even classically, since in relativity different observers would assign different points as the centre of mass.

Given the two notions of particles in QFT, one can describe two different corresponding notions of WEP. In principle, one first needs to apply SEP in order to couple the matter fields to gravity, at the fundamental level. Assuming this is completed, the motions of both the plane-wave-particles and kink particles can be derived from the combined set of Einstein's equations and matter field equations, without any appeal to any notion of WEP. In this sense, once the trajectory of the particle in the background gravitational field has been determined from the field equations, one can verify *as a theorem* whether the particle satisfies WEP or not.

Specifically, in the case of a matter field coupled to general relativity such that it locally resembles a plane wave, one can apply the WKB approximation to demonstrate that the wave 4-vector $k^\mu(x)$, orthogonal to the wavefront at its every point $x \in \mathbb{R}^4$, will satisfy a geodesic equation,

$$k^\mu(x)\nabla_\mu k^\lambda(x) = 0\,. \tag{1}$$

However, given that the plane-wave-particle is completely delocalised in space, the fact that the wave 4-vector satisfies the geodesic equation could hardly be interpreted as "the particle following a geodesic trajectory", and thus obeying WEP. Indeed, identifying the vector field orthogonal to the wavefront to the notion of "particle's trajectory" is at best an abuse of terminology.

Next, in the case of the kink particle coupled to general relativity, one assumes the configuration of the background gravitational field is such that the particle maintains its structure and that its size can be completely neglected. One can then apply the procedure given in [26,32–46] to demonstrate that the 4-vector $u^\mu(\tau)$, tangent to the kink's world line (i.e., its trajectory), will satisfy a geodesic equation ($\tau \in \mathbb{R}$ represents kink's proper time),

$$u^\mu(\tau)\nabla_\mu u^\lambda(\tau) = 0\,. \tag{2}$$

Thus, one concludes that the kink obeys WEP as a *theorem* in field theory, without the necessity to actually postulate it.

Note the crucial difference between Equations (1) and (2)—while the former features 4-dimensional variable x, the latter is given in terms of only 1-dimensional proper time τ. This reflects the fact that the plane-wave-particle is a highly delocalised object, with no well-defined position and trajectory, while the kink is a highly localised object, with a well-defined position and trajectory. As a consequence, WEP can be formulated only for the kink, and not for the plane-wave particle.

In the case of the kink, it is also important to emphasise that the zero-size approximation of the kink is crucial. Namely, without this assumption, the particle will feel the tidal forces of gravity across its size, effectively coupling its angular momentum $J^{\mu\nu}(\tau)$ to the curvature of the background gravitational field [32–46] (see also [59] for a more refined analysis of tidal effects). This will give rise to an equation of motion for the kink of the form

$$u^\mu(\tau)\nabla_\mu u^\lambda(\tau) = R^\lambda{}_{\mu\rho\sigma}u^\mu(\tau)J^{\rho\sigma}(\tau)\,, \tag{3}$$

which features explicit coupling to curvature (absent from (2)) and thus fails to obey WEP. In this sense, for realistic kink solutions WEP is *always violated*, and can be considered to hold only as an approximation when the size of the particle can be completely neglected compared to the radius of curvature of the background gravitational field. If in addition the kink has negligible total energy, it can be used as a point-like test particle.

In the above discussion, while matter fields are described as quantum, using QFT, the background gravitational field is considered to be completely classical. It should therefore not be surprising that WEP may fail to hold if one allows the gravitational field to be quantum, such as matter fields, and one needs to revisit all steps of the above analysis from the perspective of QG. In fact, the case of the kink particle has been studied in precisely this scenario [26], and it has been shown that if the background gravitational field is in a specific type of quantum superposition of different configurations, the kink will fail to obey WEP even in the zero size approximation. Simply put, the equation of motion for the kink will contain extra terms due to the interference effects between superposed configurations of gravity, giving rise to an effective force that pushes the kink off the geodesic trajectory. Moreover, of course, similar to the case of classical gravity, the resulting conclusion is a *theorem*, which follows from the fundamental field equations of the theory. One of the assumptions of that theorem is that the field equations allow for kink solutions in the first place. Namely, it is entirely possible that in quantum gravity particles cannot be localised at all, as opposed to the classical case where such an approximation can be feasible. If that is the case, then one cannot even formulate (i.e., generalise from classical theory) the notion of WEP in a quantum gravity setup. However, one can instead assume that kink solutions do exist, as was performed in [26], where a particular superposition of gravitational fields was considered, describing small quantum fluctuations around a dominant classical geometry. It was then argued that such superpositions are compatible with the approximation of a well-defined localised particle (see the discussion around Equations (2.2) and (3.15), as well

as Section 3.4 of that paper). As it turns out, even in such cases the trajectory of the kink fails to obey WEP. Therefore, the generalisations of WEP and other approximate versions of EP are not the best candidates for analysing the properties of quantum gravity.

Moreover, the assumption of a well-defined notion of a particle in the QG framework can also be supported from the point of view of nonrelativistic limit. Namely, in [4,5] an experiment was proposed in which the effects of QG fluctuations are expected to be observable, by measuring the motion of nonrelativistic particles. Furthermore, an extension of this experiment was also suggested [60], which aims to determine the potential difference between gravitational and inertial masses of those particles in such a setup. In fact, the relation between the two types of masses in the nonrelativistic limit has also been previously analysed in [26], predicting their difference due to quantum fluctuations of geometry. In this sense, the notion of a kink should make sense even in the QG setup, at least in the nonrelativistic limit.

For the case of the plane-wave-particle travelling through the superposed background of two gravitational field configurations, the analysis of the equation of motion for the wave-vector field $k^\mu(x)$, in the sense of [26,32–46], has not been performed so far (to the best of our knowledge). However, in principle, one can expect a similar interference term to appear in the WKB analysis, and give rise to a non-geodesic equation for the wave 4-vector as well. In this sense, it is to be expected that generically even the wavefronts of such plane-wave-particles would fail to obey WEP.

5. Conclusions and Discussion

In this paper, we give an overview of the equivalence principle and its various flavours formulated historically, with a special emphasis on the weak equivalence principle. We performed a critical analysis of the notions of particle and trajectory in various frameworks of physics, showing that the notion of a point-like particle and its trajectory are not always well defined. This in turn suggests that WEP might not be the best starting point for generalisation to QG, as we argue in more detail below.

As discussed in Section 4, in [26] it was shown that if superpositions of states of gravity and matter are allowed, WEP can be violated. It is important to note that the cases considered in [26] feature a specific type of superposition of three groups of states: the first consists of a single so-called dominant state—a classical state whose expectation values of the metric and the stress-energy tensors satisfy Einstein field equations; the second consists of states similar to the dominant one, with arbitrary coefficients; and the third consists of states quasi-orthogonal to the dominant one, but with negligible coefficients. Only then one may talk of a (well-localised) trajectory of the test particle in the overall superposed state and consequently about the straightforward generalisation of the classical WEP to the realm of QG. Considering that for the dominant state, being classical, the trajectory of the test particle follows the corresponding geodesic, we see that in the superposed state its trajectory would *deviate from the geodesic of the dominant state*, thus violating WEP. Note that, as discussed in Section 4, this deviation, in addition to classically weighted trajectories of the individual branches, also features purely quantum (i.e., off-diagonal) interference terms.

On the other hand, a number of recent studies propose generalisations of WEP to QG framework, arguing that it remains satisfied in such scenarios, a result *seemingly* at odds with [26]. For example, in [29–31], the authors consider superpositions of an arbitrary number of classical quasi-orthogonal states with arbitrary coefficients, arguing that since WEP is valid in each classical branch, it is valid in its superposition as well. If taken as a *definition* of what it means that a certain principle is satisfied in a superposition of different quantum states, then the above statement is manifestly true. As such, being a definition, it tells little about physics—it merely rephrases one statement ("principle A is separately satisfied in all branches of a superposition") with another, simpler ("principle A is satisfied in a superposition"). Namely, note that in [29,30], such a generalised version of EP plays no functional role in the analyses conducted in those papers. What does play a functional role is the statement of one version of classical EP (specifically, local equality between gravity

and inertia) applied to each particular state in a superposition. All physically relevant (and otherwise interesting) conclusions of the two papers could be equally obtained without ever talking about the generalised EP. In addition, in [31] EP itself is not even the main focus of the paper, and its generalisation is just introduced in analogy to the analysis of the conservation laws, which is itself an interesting topic. On the other hand, in the case of weakly superposed gravitational fields, such as in proposed experiments [4,5], the violation of the equality of inertial and gravitational masses is to be expected [26,60]. Moreover, following the spirit of the above definition, one could be misled to conclude that the notions of the particle's position and trajectory are always well-defined, as long as they are defined in each (quasi-classical) branch of the superposition.

An alternative approach to the generalisation of EP to the quantum domain was proposed in [16,20,27,28]. In those works, the authors discuss the coupling of a spatially delocalised wave-particle to gravity, with the aim of generalising such a scenario to QG. To that end, they prove a theorem which essentially states that for such a delocalised wave-particle, even when it is entangled with the gravitational field, one can always find a quantum reference frame transformation, such that in the vicinity of a given spacetime point one has a locally inertial coordinate system. The theorem employs the novel techniques of quantum reference frames (QRF) to generalise to the quantum domain the well-known result from differential geometry, that in the infinitesimal neighbourhood of any spacetime point one can always choose a locally inertial coordinate system.

The authors then employ the theorem to generalise one flavour of EP to the quantum domain. Specifically, even if the wave-particle is entangled with the gravitational field, one can use the appropriate QRF transformation to switch to a locally inertial coordinate system, and then in that system "all the (nongravitational) laws of physics must take on their familiar non-relativistic form". Here, to the best of our understanding, the phrase "non-gravitational laws of physics" refers to the equations of motion for a quantum-mechanical wave-particle, while "non-relativistic form" means that these equations of motion take the same form as in special-relativistic context.

Our understanding is that the above wave-particle generalisation of EP lies somewhere "in between" mechanics and field theory, i.e., it is in a sense stronger than WEP, which discusses particles, but weaker than SEP, which discusses full-blown matter fields. Since it refers to wave-particles rather than kinks, our analysis of WEP and its reliance on the particle trajectory does not apply to this version of EP.

The methodology in [16,20,27,28] is that one should try to generalise even approximate flavours of EP, as a stopgap result in a bigger research programme, in the hope that they may still shed some light on QG. This is of course a legitimate methodology, and from that point of view these kinds of generalisations of EP to the quantum domain represent interesting results. Nevertheless, we also believe it would be preferable to formulate a generalisation of SEP, and in a way which does not appeal to reference frames at all, since that would be closer to the essence of the statement of EP, as discussed in Section 2.

To conclude, our analysis suggests that, instead of trying to generalise various approximate formulations of EP, one should rather talk of operationally verifiable statements regarding the (in)equality of gravitational and inertial masses, possible deviation from the geodesic motion of test particles, the universality of free fall, etc., and study other principles and their possible generalisations to QG, such as SEP (see Section 4.2 in [26]), background independence, quantum nonlocality, and so on.

Author Contributions: Investigation, N.P. and M.V.; writing—original draft preparation, N.P. and M.V.; writing—review and editing, N.P. and M.V. All authors have read and agreed to the published version of the manuscript.

Funding: NP's work was partially supported by SQIG—Security and Quantum Information Group of Instituto de Telecomunicações, by Programme (COMPETE 2020) of the Portugal 2020 framework [Project Q.DOT with Nr. 039728 (POCI-01-0247-FEDER-039728)] and the Fundação para a Ciência e a Tecnologia (FCT) through national funds, by FEDER, COMPETE 2020, and by Regional Op-

erational Program of Lisbon, under UIDB/50008/2020 (actions QuRUNNER, QUESTS), Projects QuantumMining POCI-01-0145-FEDER-031826, PREDICT PTDC/CCI-CIF/29877/2017, CERN/FIS-PAR/0023/2019, QuantumPrime PTDC/EEI-TEL/8017/2020, as well as the FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT000. MV was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, and by the Science Fund of the Republic of Serbia, grant 7745968, "Quantum Gravity from Higher Gauge Theory 2021"—QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

Institutional Review Board Statement: Not applicable.

Acknowledgments: The authors wish to thank Časlav Brukner, Flaminia Giacomini, Chiara Marletto and Vlatko Vedral for useful discussions. MV is also indebted to Milovan Vasilić and Igor Salom for clarifications regarding the notion of symmetry localisation.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Here, we provide a detailed example of the two applications of the EP. First, we discuss the gravitational EP and apply it to a real scalar field, giving all mathematical details and discussing various related aspects such as locality, symmetry localisation, and so on. Then, we turn to the application of the gauge field generalisation of EP, using electrodynamics as an example. We describe how one can couple matter to an EM field, mimicking the previous gravitational example, and emphasize the analogy between the gravitational and EM case at each step. Note also that the non-Abelian gauge fields can be studied in exactly the same way. Finally, we discuss the case of test particles, and the violation of the WEP in both gravitational and electromagnetic cases.

Throughout this section, we assume that the Minkowski metric $\eta_{\mu\nu}$ has signature $(-,+,+,+)$.

Appendix A.1. The Gravitational Case

Let us begin with an example of a real scalar field in Minkowski spacetime, and apply the equivalence principle by coupling it to gravity. The equation of motion in this case is the ordinary Klein–Gordon equation,

$$\left(\eta^{\mu\nu}\partial_\mu\partial_\nu - m^2\right)\phi(x) = 0\,. \tag{A1}$$

As it stands, it describes the free scalar field in Minkowski spacetime, in an inertial coordinate system. In order to couple it to gravity (in the framework of GR), we first rewrite this equation into an arbitrary curvilinear coordinate system, as

$$\left(\tilde{g}^{\mu\nu}\tilde{\nabla}_\mu\tilde{\nabla}_\nu - m^2\right)\phi(\tilde{x}) = 0\,. \tag{A2}$$

Here the covariant derivative $\tilde{\nabla}_\mu$ is defined in terms of the Levi-Civita connection,

$$\tilde{\Gamma}^\lambda{}_{\mu\nu} = \frac{1}{2}\tilde{g}^{\lambda\sigma}\left(\partial_\mu\tilde{g}_{\nu\sigma} + \partial_\nu\tilde{g}_{\mu\sigma} - \partial_\sigma\tilde{g}_{\mu\nu}\right), \tag{A3}$$

which is in turn defined in terms of the curvilinear Minkowski metric $\tilde{g}_{\mu\nu}$. Note that the tilde symbol denotes the fact that this metric has been obtained by a coordinate transformation $\tilde{x}^\mu = \tilde{x}^\mu(x)$ from the Minkowski metric in an inertial coordinate system, $\eta_{\mu\nu}$,

$$\tilde{g}_{\mu\nu} = \frac{\partial x^\rho}{\partial\tilde{x}^\mu}\frac{\partial x^\sigma}{\partial\tilde{x}^\nu}\eta_{\rho\sigma}\,, \tag{A4}$$

and, therefore, if one were to evaluate the Riemann curvature tensor using $\tilde{g}_{\mu\nu}$ and $\tilde{\Gamma}^{\lambda}{}_{\mu\nu}$, according to the equation

$$R^{\lambda}{}_{\rho\mu\nu} = \partial_{\mu}\tilde{\Gamma}^{\lambda}{}_{\rho\nu} - \partial_{\nu}\tilde{\Gamma}^{\lambda}{}_{\rho\mu} + \tilde{\Gamma}^{\lambda}{}_{\sigma\mu}\tilde{\Gamma}^{\sigma}{}_{\rho\nu} - \tilde{\Gamma}^{\lambda}{}_{\sigma\nu}\tilde{\Gamma}^{\sigma}{}_{\rho\mu}\,, \tag{A5}$$

one would obtain that $R^{\lambda}{}_{\mu\nu\rho} = 0$ at every point in spacetime since transforming into a different coordinate system cannot induce the curvature of spacetime.

Now one can apply EP (in this example specifically SEP) in order to couple the scalar field to gravity. The statement of SEP is that, in the presence of a gravitational field (i.e., in curved spacetime), the equation of motion for the scalar field should locally retain the same form as in the absence of the gravitational field (i.e., in flat spacetime). Since Equation (A2) depends only on the field at a given spacetime point and its first and second derivatives at the same point, the equation is in fact local—it is defined within an infinitesimal neighbourhood of a single point. Given this, EP states that the corresponding equation of motion in the presence of gravity should have precisely the same form:

$$\left(g^{\mu\nu}\nabla_{\mu}\nabla_{\nu} - m^2\right)\phi(x) = 0\,. \tag{A6}$$

The absence of the tilde now denotes the fact that the covariant derivative ∇_{μ} is defined in terms of a generic Levi-Civita connection $\Gamma^{\lambda}{}_{\mu\nu}$ which is in turn defined in terms of a generic metric $g_{\mu\nu}$, which does not necessarily satisfy (A4). In other words, EP postulates that the Equation (A6) now holds even in curved spacetime since for a generic metric and connection, the Riemann curvature tensor need not be equal to zero everywhere. The interaction between the scalar field and gravity, as postulated by EP and implemented in Equation (A6), is also known in the literature as the *minimal coupling* prescription [61].

In order to convince oneself that the preparation step of transforming (A1) to (A2) is trivial in the sense that it does not introduce any substantial modification of (A1), one can additionally demonstrate that (A6) is in fact locally equivalent to (A1) as well. Namely, according to a theorem in differential geometry (see for example the end of Chapter 85 in [62]), at any specific spacetime point x_0 one can choose the locally inertial coordinate system, in which the generic metric $g_{\mu\nu}$, the corresponding connection $\Gamma^{\lambda}{}_{\mu\nu}$ and consequently also the covariant derivative ∇_{μ} take their usual Minkowski values,

$$g_{\mu\nu}(x_0) = \eta_{\mu\nu}\,, \qquad \Gamma^{\lambda}{}_{\mu\nu}(x_0) = 0\,, \qquad \nabla_{\mu}\Big|_{x=x_0} = \partial_{\mu}\,, \tag{A7}$$

so that in the infinitesimal neighbourhood of the point x_0 Equation (A6) obtains the form precisely equal to (A1).

However, note that when *integrating* (A6), one must take into account that spacetime is curved since integration is a nonlocal operation, and the locally inertial coordinate system cannot eliminate spacetime curvature. Therefore, the *solutions* of (A6) will in general be *different* from solutions of (A1), indicating the physical interaction of the scalar field with gravity, despite the fact that the form of the equation of motion is identical in both cases.

Another thing that should be emphasised is that EP itself is not a mathematical theorem, but rather a principle with physical content, since it can be either satisfied or violated. Specifically, we could have prescribed a different coupling of the scalar field to gravity, such that in curved spacetime its equation of motion takes for example the form

$$\left(g^{\mu\nu}\nabla_{\mu}\nabla_{\nu} - m^2 + R^2 + K^2\right)\phi(x) = 0\,, \tag{A8}$$

where $R \equiv R^{\mu\nu}{}_{\mu\nu}$ and $K \equiv R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}$ are the curvature scalar and Kretschmann invariant, respectively. This equation is not equivalent to (A2) and there is no coordinate system in which it can take the form (A1) since R and K are invariants. In this sense, (A8) is an example of a scalar field coupled to gravity such that EP is violated. This type of interaction between matter and gravity is also known in the literature as *non-minimal coupling* [61].

Finally, we should note that the transformation from (A1) to (A2) amounts to what is also known in the literature as *symmetry localisation* [61]. In particular, one can verify that (A1) remains invariant with respect to the group $\mathbb{R}^4$ of global translations,

$$x^\mu \to \tilde{x}^\mu = x^\mu + \zeta^\mu\,, \qquad (\zeta \in \mathbb{R}^4)\,, \tag{A9}$$

while (A2) remains invariant with respect to the group $Diff(\mathbb{R}^4)$ of spacetime diffeomorphisms, obtained by localisation of the translational symmetry group,

$$x^\mu \to \tilde{x}^\mu = x^\mu + \zeta^\mu(x) \equiv \tilde{x}^\mu(x)\,, \tag{A10}$$

which represent general curvilinear coordinate transformations, used in (A4). One can explicitly verify that all three Equations (A2), (A6) and (A8) remain invariant with respect to local translations (A10) while describing no coupling to gravity, coupling to gravity that satisfies EP, and coupling to gravity that violates EP, respectively. In this sense, contrary to a common misconception (often stated in the literature) that symmetry localisation gives rise to interactions, one can say that the process of symmetry localisation *does not* introduce nor prescribe interactions in any way whatsoever. In particular, a direct counterexample is the Equation (A4), which manifestly *does* obey local translational symmetry, while it *does not* give rise to any interaction whatsoever (see below for the analogous counterexample in electrodynamics).

Appendix A.2. The Electromagnetic Case

Let us begin with an example of a Dirac field in Minkowski spacetime, and apply the generalised equivalence principle by coupling it to the EM field. The equation of motion in this case is the ordinary Dirac equation,

$$\left(i\gamma^\mu\partial_\mu - m\right)\psi(x) = 0\,, \tag{A11}$$

where γ^μ are standard Dirac gamma matrices, satisfying the anticommutator identity of the Clifford algebra $\{\gamma^\mu, \gamma^\nu\} = -2\eta^{\mu\nu}$. As it stands, Equation (A11) describes the free Dirac field, not coupled to an EM field in any way. Note that it is invariant with respect to global $U(1)$ transformations, defined as

$$\psi \to \psi' = e^{-i\lambda}\psi\,, \qquad e^{-i\lambda} \in U(1)\,, \qquad \lambda \in \mathbb{R}\,. \tag{A12}$$

In order to couple it to standard Maxwell electrodynamics, we first rewrite this equation into a form which is invariant with respect to local $U(1)$ transformations,

$$\psi \to \psi' = e^{-i\lambda(x)}\psi\,, \qquad \partial_\mu \to \tilde{\mathcal{D}}_\mu = \partial_\mu + i\partial_\mu\lambda(x)\,, \tag{A13}$$

so that the equation takes the form

$$\left(i\gamma^\mu\tilde{\mathcal{D}}_\mu - m\right)\psi(x) = 0\,, \tag{A14}$$

Note that here, $\tilde{\mathcal{D}}$ denotes the covariant derivative with respect to the "pure gauge" connection

$$\tilde{A}_\mu \equiv \partial_\mu\lambda(x)\,, \tag{A15}$$

where $\lambda(x)$ denotes the arbitrary gauge function. Moreover, note that (A11) is analogous to (A1), (A14) is analogous to (A2), while the global and local $U(1)$ gauge transformations (A12) and (A13) are EM analogues of the global and local spacetime translations (A9) and (A10) from the gravitational case. Finally, note that if one were to evaluate the electromagnetic Faraday field strength tensor using $\tilde{A}_\mu$ from (A15), according to the equation

$$F_{\mu\nu} = \partial_\mu\tilde{A}_\nu - \partial_\nu\tilde{A}_\mu\,, \tag{A16}$$

one would obtain that $F_{\mu\nu} = 0$ at every point in spacetime since the potential that is a pure gauge cannot induce an EM field. Here (A16) is analogous to (A5).

Once the Dirac equation in the form (A14) is in hand, one can apply the electromagnetic generalisation of EP in order to couple the Dirac field to an EM field. The statement of EP, in this case, is that in the presence of an EM field, the equation of motion for the Dirac field should locally retain the same form as in the absence of the EM field. Since Equation (A14) depends only on the field at a given spacetime point and its first derivatives at the same point, it is therefore defined within an infinitesimal neighbourhood of a single point—in other words, it is local. Given this, electromagnetic EP states that the corresponding equation of motion in the presence of EM field should have precisely the same form (the analogue of (A6)):

$$\left(i\gamma^{\mu}\mathcal{D}_{\mu} - m\right)\psi(x) = 0\,. \tag{A17}$$

The absence of the tilde now denotes the fact that the covariant derivative $\mathcal{D}_{\mu} \equiv \partial_{\mu} + iA_{\mu}$ is defined in terms of a generic $U(1)$ connection A_{μ} which does not necessarily satisfy (A15), but does obey the usual gauge transformation rule,

$$A_{\mu} \to A'_{\mu} = A_{\mu} + \partial_{\mu}\lambda(x)\,. \tag{A18}$$

In other words, electromagnetic EP postulates that the Equation (A17) holds even in the presence of an EM field since for a generic connection A_{μ} the Faraday tensor may not be equal to zero everywhere. The interaction between the Dirac field and the EM field as postulated by the electromagnetic EP and implemented in Equation (A17) is again known in the literature as the *minimal coupling* prescription [61,63].

If one wishes to convince oneself that the preparation step of transforming (A11) to (A14) is trivial in the sense that it does not introduce any substantial modification of (A11), one can additionally demonstrate that (A17) is in fact locally equivalent to (A11). To do this, at any specific spacetime point x_0 one can choose the following $U(1)$ gauge,

$$\lambda(x) = -A_{\mu}(x_0)x^{\mu}\,, \tag{A19}$$

so that, according to (A18)

$$A'_{\mu}(x) = A_{\mu}(x) - \partial_{\mu}(A_{\nu}(x_0)x^{\nu}) \qquad \Rightarrow \qquad A'_{\mu}(x_0) = 0\,, \qquad \mathcal{D}_{\mu}\Big|_{x=x_0} = \partial_{\mu}\,. \tag{A20}$$

This choice of gauge is the EM analogue of the choice of a locally inertial coordinate system (A7). Substituting this into the primed version of (A17) and evaluating the whole equation at $x = x_0$, it reduces precisely to the form (A11) in the infinitesimal neighbourhood at that point, despite the presence of nonzero EM field.

Again note that when *integrating* (A17), one must take into account that the EM field is nonzero since integration is a nonlocal operation, and the choice of gauge (A19) eliminates the EM potential from (A17) only at the point x_0, while the Faraday tensor is gauge invariant. Therefore, the *solutions* of (A17) will in general be *different* from solutions of (A11), indicating the physical interaction of the Dirac field with EM field, despite the fact that the form of the equation of motion for the Dirac field is identical in both cases.

As in the case of gravity, we should emphasise that the electromagnetic EP is not a mathematical theorem, but rather a principle with physical content, since it can be either satisfied of violated. Specifically, we could have prescribed a different coupling of the Dirac field to electrodynamics, such that in the presence of an EM field its equation of motion takes for example the form (analogue of (A8))

$$\left(i\gamma^{\mu}\mathcal{D}_{\mu} - m + I_1 + I_2\right)\psi(x) = 0\,, \tag{A21}$$

where $I_1 \equiv F^{\mu\nu}F_{\mu\nu}$ and $I_2 \equiv \varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$ are the two fundamental invariants of the Faraday tensor. This equation is not equivalent to (A14), and there exists no local $U(1)$ gauge in which it could take the form (A11), since I_1 and I_2 are invariants. In this sense, (A21) is

an example of a Dirac field coupled to the EM field such that the electromagnetic EP is violated. This is also known in the literature as *non-minimal coupling* [61,63].

Finally, we should also note that the transformation from (A11) to (A14) amounts to what is also known in the literature as *symmetry localisation* [61,63]. Specifically, one can explicitly verify that all three Equations (A14), (A17) and (A21) remain invariant with respect to local $U(1)$ gauge transformations, while describing no coupling to an EM field, coupling to an EM field that satisfies the electromagnetic EP, and coupling to an EM field that violates electromagnetic EP, respectively. In this sense, one can again say that the process of symmetry localisation *does not* introduce nor prescribe interactions in any way whatsoever. In the case of electrodynamics and other gauge theories, this is quite often misrepresented in literature—the step of symmetry localisation is silently joined together with the step of applying the electromagnetic version of EP; thus, in the end, giving rise to an interacting theory, and the resulting presence of the interaction is then mistakenly attributed to the localisation of symmetry, rather than to the application of EP. Similar to the gravitational case above, the equation of motion (A14) is an explicit counterexample to such an attribution, since it *does* have local $U(1)$ symmetry, but *does not* have any interaction with an EM field.

Appendix A.3. The Test Particle Case

The last topic we should address is the context in which the statement of electromagnetic EP is compatible with the existence of the Lorentz force law, acting on charged test particles. Namely, one often distinguishes the motion of a test particle in a gravitational field from a motion of a test particle in an EM field, by comparing the geodesic Equation (2)

$$u^{\mu}(\tau)\nabla_{\mu}u^{\lambda}(\tau) = 0\,, \tag{A22}$$

where u^{μ} is the 4-velocity of the test particle, with the Lorentz force equation

$$u^{\mu}(\tau)\nabla_{\mu}u^{\lambda}(\tau) = \frac{q}{m}F^{\lambda\rho}\,u_{\rho}(\tau)\,, \tag{A23}$$

where q/m is the charge-to-mass ratio of a test particle moving in an external EM field, described by the Faraday tensor $F_{\mu\nu}$. A typical conclusion one draws from this comparison is that the interaction with the EM field gives rise to a "real force", while the interaction with the gravitational field does not.

However, it is highly misleading to compare (A22) to (A23) in the first place. Namely, as we have discussed in detail in Section 4, in field theory the notion of a particle can be defined only approximately, and this applies equally for electrodynamics as well as for gravity. Specifically, given the example discussed above, of a Dirac field coupled to an EM field via Equation (A17), we have seen that in the infinitesimal neighbourhood of a given point x_0 one can completely gauge away any presence of the coupling to EM field from (A17). In this sense, the notion of a test particle that satisfies (A23) cannot be identified with an idealised point-particle, that has exactly zero size. Instead, the realistic test particle is a wave-packet configuration of a Dirac field (a kink), and as such has a small but nonzero size. As it evolves, the different parts of the wave-packet are subject to interaction with the EM potential A_{μ} at *different* points of spacetime, giving rise to an effective non-minimal coupling with the Faraday tensor $F_{\mu\nu}$. This is completely analogous to the case of a test particle with small but nonzero size interacting with spacetime curvature due to tidal forces—both effects are equally nonlocal since both kinks have nonzero size. On the other hand, a test particle that satisfies (A22) represents an idealised point-particle (a leading order approximation in the multipole expansion of the matter field), i.e., a kink which thus has precisely zero size.

In this sense, the Lorentz force Equation (A23) rather ought to be compared with the Papapetrou Equation (3),

$$u^{\mu}(\tau)\nabla_{\mu}u^{\lambda}(\tau) = R^{\lambda}{}_{\mu\rho\sigma}\,u^{\mu}(\tau)J^{\rho\sigma}(\tau)\,. \tag{A24}$$

Indeed, one can see quite a reasonable analogy between (A23) and (A24). There are of course small technical differences due to the precise nature of the coupling to various moments of the kink, but nevertheless, the two equations are strikingly similar. Given this, while one can still draw the conclusion that the interaction of a kink with the EM field gives rise to a "real force", one can draw precisely the same conclusion for the interaction of a kink with the gravitational field. There is no distinction between gravity and the other gauge interactions at this level—all four interactions in nature (strong, weak, electromagnetic and gravitational) are equally "real". In addition, all four interactions satisfy EP at the fundamental field theory level (i.e., in the sense of strong generalised EP), while at the level of mechanics, a corresponding weak generalised EP is manifestly violated in all four cases.

References

1. Kay, B.S. Decoherence of macroscopic closed systems within Newtonian quantum gravity. *Class. Quantum Gravity* **1998**, *15*, L89. [CrossRef]
2. Oniga, T.; Wang, C.H.T. Quantum gravitational decoherence of light and matter. *Phys. Rev. D* **2016**, *93*, 044027. [CrossRef]
3. Bruschi, D.E. On the weight of entanglement. *Phys. Lett. B* **2016**, *754*, 182. [CrossRef]
4. Bose, S.; Mazumdar, A.; Morley, G.W.; Ulbricht, H.; Toroš, M.; Paternostro, M.; Geraci, A.A.; Barker, P.F.; Kim, M.S.; Milburn, G. Spin entanglement witness for quantum gravity. *Phys. Rev. Lett.* **2017**, *119*, 240401. [CrossRef]
5. Marletto, C.; Vedral, V. Gravitationally Induced Entanglement between Two Massive Particles is Sufficient Evidence of Quantum Effects in Gravity. *Phys. Rev. Lett.* **2017**, *119*, 240402.
6. Marletto, C.; Vedral, V. When can gravity path-entangle two spatially superposed masses? *Phys. Rev. D* **2018**, *98*, 046001. [CrossRef]
7. Paunković, N.; Vojinović, M. Gauge protected entanglement between gravity and matter. *Class. Quantum Gravity* **2018**, *35*, 185015. [CrossRef]
8. Oreshkov, O.; Costa, F.; Brukner, Č. Quantum correlations with no causal order. *Nat. Commun.* **2012**, *3*, 1092. [CrossRef]
9. Araújo, M.; Branciard, C.; Costa, F.; Feix, A.; Giarmatzi, C.; Brukner, Č. Witnessing causal nonseparability. *New J. Phys.* **2015**, *17*, 102001. [CrossRef]
10. Vilasini, V. An Introduction to Causality in Quantum Theory (and Beyond). Master's Thesis, ETH, Zürich, Switzerland, 2017.
11. Oreshkov, O. Time-delocalized quantum subsystems and operations: On the existence of processes with indefinite causal structure in quantum mechanics. *Quantum* **2019**, *3*, 206. [CrossRef]
12. Paunković, N.; Vojinović, M. Causal orders, quantum circuits and spacetime: Distinguishing between definite and superposed causal orders. *Quantum* **2020**, *4*, 275. [CrossRef]
13. Vilasini, V.; Colbeck, R. General framework for cyclic and fine-tuned causal models and their compatibility with space-time. *Phys. Rev. A* **2022**, *106*, 032204. [CrossRef]
14. Vilasini, V.; Renner, R. Embedding cyclic causal structures in acyclic spacetimes: No-go results for process matrices. *arXiv* **2022**, arXiv:2203.11245.
15. Ormrod, N.; Vanrietvelde, A.; Barrett, J. Causal structure in the presence of sectorial constraints, with application to the quantum switch. *arXiv* **2022**, arXiv:2204.10273.
16. Giacomini, F.; Castro-Ruiz, E.; Brukner, Č. Quantum mechanics and the covariance of physical laws in quantum reference frames. *Nat. Commun.* **2019**, *10*, 494. [CrossRef] [PubMed]
17. Vanrietvelde, A.; Höhn, P.A.; Giacomini, F.; Castro-Ruiz, E. A change of perspective: Switching quantum reference frames via a perspective-neutral framework. *Quantum* **2020**, *4*, 225. [CrossRef]
18. Krumm, M.; Höhn, P.A.; Müller, M.P. Quantum reference frame transformations as symmetries and the paradox of the third particle. *Quantum* **2021**, *5*, 530. [CrossRef]
19. Ahmad, S.A.; Galley, T.D.; Höhn, P.A.; Lock, M.P.E.; Smith, A.R.H. Quantum Relativity of Subsystems. *Phys. Rev. Lett.* **2022**, *128*, 170401. [CrossRef]
20. de la Hamette, A.C.; Kabel, V.; Castro-Ruiz, E.; Brukner, Č. Falling through masses in superposition: Quantum reference frames for indefinite metrics. *arXiv* **2021**, arXiv:2112.11473.
21. Colladay, D.; Kostelecký, V.A. Lorentz-violating extension of the standard model. *Phys. Rev. D* **1998**, *58*, 116002. [CrossRef]
22. Kostelecký, V.A.; Russell, N. Data tables for Lorentz and *CPT* violation. *Rev. Mod. Phys.* **2011**, *83*, 11–31.
23. Amelino-Camelia, G. Particle-dependent deformations of Lorentz symmetry. *Symmetry* **2012**, *4*, 344–378. [CrossRef]
24. Amelino-Camelia, G.; Palmisano, M.; Ronco, M.; D'Amico, G. Mixing coproducts for theories with particle-dependent relativistic properties. *Int. J. Mod. Phys. D* **2020**, *29*, 2050017. [CrossRef]
25. Torri, M.D.C.; Antonelli, V.; Miramonti, L. Homogeneously Modified Special relativity (HMSR). *Eur. Phys. J. C* **2019**, *79*, 808.
26. Pipa, F.; Paunković, N.; Vojinović, M. Entanglement-induced deviation from the geodesic motion in quantum gravity. *J. Cosmol. Astropart. Phys.* **2019**, *2019*, 57. [CrossRef]
27. Giacomini, F.; Brukner, Č. Einstein's Equivalence principle for superpositions of gravitational fields. *arXiv* **2020**, arXiv:2012.13754.

28. Giacomini, F.; Brukner, Č. Quantum superposition of spacetimes obeys Einstein's equivalence principle. *AVS Quantum Sci.* **2022**, *4*, 015601.
29. Marletto, C.; Vedral, V. On the testability of the equivalence principle as a gauge principle detecting the gravitational t^3 phase. *Front. Phys.* **2020**, *8*, 176.
30. Marletto, C.; Vedral, V. Sagnac interferometer and the quantum nature of gravity. *J. Phys. Commun.* **2021**, *5*, 051001.
31. Marletto, C.; Vedral, V. The quantum totalitarian property and exact symmetries. *AVS Quantum Sci.* **2022**, *4*, 015603. [CrossRef]
32. Einstein, A.; Infeld, L.; Hoffmann, B. The Gravitational Equations and the Problem of Motion. *Ann. Math.* **1938**, *39*, 65. [CrossRef]
33. Mathisson, M. Neue mechanik materieller systemes. *Acta Phys. Pol.* **1937**, *6*, 163.
34. Papapetrou, A. Spinning test-particles in general relativity, I. *Proc. R. Soc. A* **1951**, *209*, 248.
35. Tulczyjev, W. Equations of motion of rotating bodies in general relativity theory. *Acta Phys. Pol.* **1959**, *18*, 393.
36. Taub, A.H. Motion of Test Bodies in General Relativity. *J. Math. Phys.* **1964**, *5*, 112. [CrossRef]
37. Dixon, G. A covariant multipole formalism for extended test bodies in general relativity. *Nuovo Cim.* **1964**, *34*, 317. [CrossRef]
38. Dixon, G. Classical theory of charged particles with spin and the classical limit of the Dirac equation. *Nuovo Cim.* **1965**, *38*, 1616. [CrossRef]
39. Dixon, G. Dynamics of extended bodies in general relativity. I. Momentum and angular momentum. *Proc. R. Soc. A* **1970**, *314*, 499.
40. Dixon, G. Dynamics of extended bodies in general relativity - II. Moments of the charge-current vector. *Proc. R. Soc. A* **1970**, *319*, 509.
41. Dixon, G. The definition of multipole moments for extended bodies. *Gen. Relativ. Gravit.* **1973**, *4*, 199. [CrossRef]
42. Yasskin, P.B.; Stoeger, W.R. Propagation equations for test bodies with spin and rotation in theories of gravity with torsion. *Phys. Rev. D* **1980**, *21*, 2081. [CrossRef]
43. Nomura, K.; Shirafuji, T.; Hayashi, K. Spinning Test Particles in Spacetime with Torsion. *Prog. Theor. Phys.* **1991**, *86*, 1239. [CrossRef]
44. Nomura, K.; Shirafuji, T.; Hayashi, K. Semiclassical particles with arbitrary spin in the Riemann-Cartan space-time. *Prog. Theor. Phys.* **1992**, *87*, 1275. [CrossRef]
45. Vasilić, M.; Vojinović, M. Classical spinning branes in curved backgrounds. *JHEP* **2007**, *7*, 28. [CrossRef]
46. Vasilić, M.; Vojinović, M. Spinning branes in Riemann-Cartan spacetime. *Phys. Rev. D* **2008**, *78*, 104002. [CrossRef]
47. Accioly, A.; Paszko, R. Conflict between the Classical Equivalence Principle and Quantum Mechanics. *Adv. Stud. Theor. Phys.* **2009**, *3*, 65.
48. Longhi, S. Equivalence principle and quantum mechanics: Quantum simulation with entangled photons. *Opt. Lett.* **2018**, *43*, 226. [CrossRef]
49. Chowdhury, P.; Home, D.; Majumdar, A.S.; Mousavi, S.V.; Mozaffari, M.R.; Sinha, S. Strong quantum violation of the gravitational weak equivalence principle by a non-Gaussian wave packet. *Class. Quantum Gravity* **2012**, *29*, 025010. [CrossRef]
50. Rosi, G.; D'Amico, G.; Cacciapuoti, L.; Sorrentino, F.; Prevedelli, M.; Zych, M.; Brukner, Č.; Tino, G.M. Quantum test of the equivalence principle for atoms in coherent superposition of internal energy states. *Nat. Commun.* **2017**, *8*, 15529. [CrossRef]
51. Zych, M.; Brukner, Č. Quantum formulation of the Einstein Equivalence Principle. *Nat. Phys.* **2018**, *14*, 1027. [CrossRef]
52. Anastopoulos, C.; Hu, B.L. Equivalence principle for quantum systems: Dephasing and phase shift of free-falling particles. *Class. Quantum Gravity* **2018**, *35*, 035011. [CrossRef]
53. Hardy, L. Implementation of the Quantum Equivalence Principle. In *Progress and Visions in Quantum Theory in View of Gravity*; Finster, F., Giulini, D., Kleiner, J., Tolksdorf, J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 189–220.
54. Kretschmann, E. Über den physikalischen Sinn der Relativitätspostulate, A. Einsteins neue und seine ursprüngliche Relativitätstheorie. *Ann. Phys.* **1918**, *358*, 575–614.
55. Misner, C.W.; Thorne, K.S.; Wheeler, J.A. *Gravitation*; W. H. Freeman and Co.: San Francisco, CA, USA, 1973.
56. Okon, E.; Callender, C. Does Quantum Mechanics Clash with the Equivalence Principle—And Does it Matter? *Eur. J. Phil. Sci.* **2011**, *1*, 133. [CrossRef]
57. Casola, E.D.; Liberati, S.; Sonego, S. Nonequivalence of equivalence principles. *Am. J. Phys.* **2015**, *83*, 39. [CrossRef]
58. Viola, L.; Onofrio, R. Testing the equivalence principle through freely falling quantum objects. *Phys. Rev. D* **1997**, *55*, 455–462. [CrossRef]
59. Plyatsko, R. Gravitational ultrarelativistic spin-orbit interaction and the weak equivalence principle. *Phys. Rev. D* **1998**, *58*, 084031.
60. Bose, S.; Mazumdar, A.; Schut, M.; Toroš, M. Entanglement witness for the weak equivalence principle. *arXiv* **2022**, arXiv:2203.11628.
61. Blagojević, M. *Gravitation and Gauge Symmetries*; Institute of Physics Publishing: Bristol, UK, 2002.
62. Landau, L.D.; Lifshitz, E.M. *The Classical Theory of Fields*, 4th ed.; Butterworth-Heinemann: Oxford, UK, 1980.
63. Peskin, M.E.; Schroeder, D.V. *An Introduction to Quantum Field Theory*; Addison-Wesley Publishing Co.: Boston, MA, USA, 1995.

Journal of Physics:
Condensed Matter

PURPOSE-LED
PUBLISHING™

PAPER

Information geometric analysis of long range topological superconductors

To cite this article: Syed Tahir Amin *et al* 2019 *J. Phys.: Condens. Matter* **31** 485402

View the article online for updates and enhancements.

You may also like

- Uhlmann fidelity and fidelity susceptibility for integrable spin chains at finite temperature: exact results
Micha Biaoczyk, Fernando Javier Gómez-Ruiz and Adolfo del Campo
- Haldane model at finite temperature
Luca Leonforte, Davide Valenti, Bernardo Spagnolo et al.
- Device-independent quantum reading and noise-assisted quantum transmitters
W Roga, D Buono and F Illuminati

This content was downloaded from IP address 147.91.81.84 on 19/09/2025 at 14:14

IOP Publishing Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **31** (2019) 485402 (8pp) https://doi.org/10.1088/1361-648X/ab3c70

Information geometric analysis of long range topological superconductors

Syed Tahir Amin[1,2,4], Bruno Mera[1,3,5], Nikola Paunković[1,3] and Vítor R Vieira[2,4]

[1] Instituto de Telecomunicações, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal
[2] Departamento de Física, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
[3] Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
[4] CeFEMA, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

E-mail: bruno.mera@tecnico.ulisboa.pt

Received 8 April 2019, revised 2 August 2019
Accepted for publication 19 August 2019
Published 6 September 2019

Abstract

The Uhlmann connection is a mixed state generalisation of the Berry connection. The latter has a very important role in the study of topological phases at zero temperature. Closely related, the quantum fidelity is an information theoretical quantity which is a measure of distinguishability of quantum states. Moreover, it has been extensively used in the analysis of quantum phase transitions. In this work, we study topological phase transitions in 1D and 2D topological superconductors with long-range hopping and pairing amplitudes, using the fidelity and the quantity Δ closely related to the Uhlmann connection. The drop in the fidelity and the departure of Δ from zero signal the topological phase transitions in the models considered. The analysis of the ground state fidelity susceptibility and its associated critical exponents are also applied to the study of the aforementioned topological phase transitions.

Keywords: topological phases, phase transitions, information geometry, long range superconductivity, fidelity

(Some figures may appear in colour only in the online journal)

1. Introduction

The last decade saw a remarkable development in the research in topological quantum matter because of its unique properties which are in general immune to small local perturbations. Topological phases (TPs) of matter, like topological insulators and superconductors, are described in the bulk by topological invariants, such as the winding number of a map or the Chern number of a vector bundle over the Brillouin zone [1], which justify the robustness to perturbations preserving the gap. By changing the parameters of the Hamiltonian, such as hopping amplitudes or external fields, we can control the phase of the system by closing the gap. When this gap closing is accompanied by a change in the values of topological invariants we say that we have a topological phase transition. Such phase transitions are witnessed by using tools from quantum information such as entanglement [2–5] and, more recently, the ground state fidelity and the associated information geometric notions [6–11]. The bulk-to-boundary principle states that, if we have a system composed of two parts, described in their bulks by different values of a topological invariant, there will exist gapless excitations localised at their boundary. These so-called edge states are topologically protected.

In [12], Kitaev introduced a 1D model of a topological superconductor, with hopping and pairing amplitudes between nearest neighbours; where topological phase transitions were mainly controlled by the chemical potential μ. Since then, a number of modifications with long-range hopping and pairing amplitudes have been considered [13–23]. It was suggested that a long-range superconducting Hamiltonian may be achieved by

[5] Author to whom any correspondence should be addressed.

 © 2019 IOP Publishing Ltd Printed in the UK

using a quantum simulator such as an optical lattice with cold atoms, where changing the depth of lattice potential can mimic the exponential decaying hopping amplitudes [24]. There may be the possibility of realising such Hamiltonians in the Shiba chain [25, 26]. In the case of Kitaev's model the topological non-trivial phase occurs for $\mu \in (-1, 1)$. However, as the long-range amplitudes are switched on, the topological non-trivial phase starts to expand in phase diagram, as shown in figure 1 of [23]. On the other hand, in the case of 2D p-wave superconductors with long-range amplitudes, the TPs start to either expand or shrink, as shown in figure 1 of [24]. The fidelity and the related quantities, such as the ground state fidelity susceptibility and the Uhlmann quantity, are shown to be sensitive and to detect topological phase transitions [27–29] and the traditional symmetry breaking phase transitions [30]. Thus, we want to extend the use of above mentioned quantities to probe topological phase transitions in the long-range models given in [23], for 1D, and [24] for the 2D case.

Our paper proceeds as follows. In section 2, we briefly present the expressions for fidelity, the Uhlmann quantity Δ and the ground state fidelity susceptibility, and explain how these quantities are used to probe topological phase transitions. In section 3, we present the results of the fidelity and Uhlmann connection analysis along with the ground state fidelity susceptibility for 1D and 2D TSCs with long-range amplitudes. We also analyse the critical behaviour at zero temperature by looking at critical exponents. Finally, in section 4, we present our conclusions.

2. Information geometry and phase transitions

In this section, we present the main quantities through which topological phase transitions are studied. The fidelity quantifies the distinguishability between two mixed states ρ and σ and is given by

$$F(\rho, \sigma) = \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}. \tag{1}$$

Consider a smooth family of Hamiltonians $\{\mathcal{H}(q)\}_{q\in\mathcal{Q}}$ where $\mathcal{Q}$ is a configuration space of parameters, assumed to be a smooth manifold. For a given temperature T, we can consider the associated family of thermal states, in units where $k_B \equiv 1$,

$$\rho(q) = \frac{\exp\left(-\frac{\mathcal{H}(q)}{T}\right)}{Z(q)},\ q \in \mathcal{Q}, \tag{2}$$

where $Z(q) = \mathrm{Tr}\left[\exp\left(-\mathcal{H}(q)/T\right)\right]$. One can then consider the fidelity between two close points q and $q' = q + \delta q$. When the two points are within the same phase, the corresponding states $\rho(q)$ and $\rho(q + \delta q)$ will be almost the same and the fidelity will be close to 1. However, whenever there is a phase transition, the state of the system changes dramatically and the fidelity will drop, signalling the phase transition. In the *thermodynamic limit*, the fidelity between states from two different phases becomes zero. Moreover, generically, in the thermodynamic limit any two distinct states become orthogonal, even when taken from the same phase (consequence of the famous Anderson orthogonality catastrophe [31]). This is why we consider the finite size case, in which the fidelity is not zero since it is a product of a finite number of generically non-zero terms. Note though that this is not necessarily so. Indeed, it is known that for different phases of topological band insulators the corresponding ground state fidelity vanishes exactly [32] (see [33] for further study). In the thermodynamic limit, the quantity to be considered is the *fidelity susceptibility*, which is finite in the gapped phases, see equation (6) below, for the systems considered.

In our study, we are interested in zero temperature quantum phase transitions. Thus, one might analyse the pure state fidelity in terms of the ground states $|\psi(q)\rangle$, given by

$$F(|\psi(q)\rangle, |\psi(q+\delta q)\rangle) = |\langle\psi(q)|\psi(q+\delta q)\rangle|. \tag{3}$$

For the systems considered, the ground state manifold is non-degenerate outside the set of gapless critical points. One could thus use the above pure state expression to evaluate the ground state fidelity. Nevertheless, since in our case even for finite system size the ground state becomes degenerate at the gapless points, the pure state fidelity can be probed only *in the vicinity* of the critical points. To evaluate the fidelity at the critical points as well, we thus consider the zero temperature limit of the above mixed state expression (1). Note that at the critical points, the zero temperature limit of the thermal state is not pure, but the (normalised) projector onto the *degenerate* ground state manifold.

The notion of state distinguishability quantified by the fidelity provides associated information geometric quantities, namely, a distance and a Riemannian metric in the space of states and a connection in the fibre bundle of purifications [34]. Consider, for simplicity, a finite dimensional Hilbert space H which we can identify with $\mathbb{C}^n$. For a fixed rank k, the space of density operators is a manifold. The so-called Bures distance in this space is defined by $d_B^2(\rho, \sigma) = \big(1 - F(\rho, \sigma)\big)$. Its infinitesimal version, $d_B^2(\rho, \rho + \delta\rho)$ to second order in $\delta\rho$, defines a Riemannian metric. By taking the pullback with respect to the map $q \mapsto \rho(q)$, we obtain the fidelity susceptibility $\chi_F(q)$, which is a positive semidefinite rank-2 covariant tensor in $\mathcal{Q}$. In this work, we are considering TSCs at $T = 0$, in dimensions $D = 1, 2$, whose Hamiltonian can be expressed as

$$\begin{aligned}\mathcal{H}(q) &= \frac{1}{2}\int_{\mathrm{B.Z.}} \frac{\mathrm{d}^D k}{(2\pi)^D} \widetilde{\psi}^\dagger(\mathbf{k}) H_{\mathrm{BdG}}(\mathbf{k}; q) \widetilde{\psi}(\mathbf{k}) \\ &= \frac{1}{2}\int_{\mathrm{B.Z.}} \frac{\mathrm{d}^D k}{(2\pi)^D} \widetilde{\psi}^\dagger(\mathbf{k}) d^\mu(\mathbf{k}; q) \sigma_\mu \widetilde{\psi}(\mathbf{k}),\end{aligned} \tag{4}$$

where $\widetilde{\psi}^\dagger(\mathbf{k}) = (c^\dagger(\mathbf{k})\ c(-\mathbf{k}))$, in which $c^\dagger(\mathbf{k})$ ($c(\mathbf{k})$) creates (destroys) an electron with momentum $\mathbf{k} \in$ B.Z., where B.Z. denotes the first Brillouin zone, $\{\sigma_\mu\}_{\mu=1}^3$ are the standard Pauli matrices and we employ the Einstein summation convention. The topological phase, for fixed $q \in \mathcal{Q}$, is specified by the homotopy class of a map Φ_q : B.Z. $\to S^D$ given by

$$\Phi_q : \mathbf{k} \mapsto n^\mu(\mathbf{k}; q) = \frac{d^\mu(\mathbf{k}; q)}{|d(\mathbf{k}; q)|}, \tag{5}$$

where in the case $D = 1$, due to symmetry considerations, the target space is the equator of the Bloch sphere, therefore, a

circle S^1. The norm of the vector $d^\mu(\mathbf{k};q)$ is the gap function, $|d(\mathbf{k};q)| \equiv E(\mathbf{k};q)$. When the gap closes, the map Φ_q stops being well-defined. The expression of the ground state fidelity susceptibility is given by [29]

$$\chi_F = \chi_{\alpha\beta}(q)\mathrm{d}q^\alpha \mathrm{d}q^\beta = \frac{1}{4}\int_{\mathrm{B.Z.}} \frac{\mathrm{d}^D k}{(2\pi)^D}\delta_{\mu\nu}\frac{\partial n^\mu}{\partial q^\alpha}\frac{\partial n^\nu}{\partial q^\beta}\mathrm{d}q^\alpha \mathrm{d}q^\beta. \tag{6}$$

Note that the above thermodynamic limit expression for the fidelity susceptibility is not the zero temperature limit of the pullback of the Bures metric. Instead, it is the pullback of the Fubini–Study metric, i.e. the natural Riemannian metric in the space of pure states. This is done because the thermodynamic limit of the former, for quadratic Hamiltonians, does not see the phase transitions [29].

In addition, one can consider purifications of density matrices. A purification of a density matrix ρ is as a $n \times k$ matrix w, such that $\rho = ww^\dagger$. There is a $\mathrm{U}(k)$ gauge degree of freedom in choosing w, since wU, with $U \in \mathrm{U}(k)$, defines the same density matrix ρ. This gauge ambiguity defines the aforementioned (principal) fibre bundle of purifications over the space of density matrices of rank k. The fidelity induces a connection in this principal bundle known as the Uhlmann connection [35].

To probe the Uhlman connection, we study the quantity Δ defined by

$$\begin{aligned}\Delta(\rho(q),\rho(q+\delta q)) :&= F(\rho(q),\rho(q+\delta q)) \\ &- \mathrm{Tr}(\sqrt{\rho(q)}\sqrt{\rho(q+\delta q)}).\end{aligned} \tag{7}$$

One can show that,

$$\Delta(\rho(q),\rho(q+\delta q)) = \mathrm{Tr}\{|\sqrt{\rho(q+\delta q)}\sqrt{\rho(q)}|(I-V)\} \tag{8}$$

where V approximates the infinitesimal parallel transport according to the Uhlmann connection in the gauge specified by $w(q) = \sqrt{\rho(q)}$. Note that, when $V = I$ then Δ vanishes and, thus, this quantity is probing the non-triviality of the Uhlmann connection. Moreover, whenever the states $\rho(q)$ and $\rho(q+\delta q)$ commute we have $\Delta = 0$. In other words, the non-triviality of V, and, thus, of the Uhlmann connection, quantifies the change in the eigenbasis of $\rho(q)$. The quantity Δ quantifies part of the change of a density operator and, therefore, much like fidelity, it probes phase transitions (for more details see [27, 29, 30]). Note though that the fidelity captures the total change in the density operator, i.e. the change in *both* the eigenvalues and the eigenvectors of the density operator. Therefore, the fidelity is sensitive to a broader range of variations. Nevertheless, Δ provides finer information concerning state distinguishability. The analysis of both quantities thus allows us to probe what kind of variations of the density operator lead to phase transitions.

The ground state of the topological phase is obtained by filling the occupied bands below the gap and it is non-degenerate. For the systems considered here, there is a single occupied band and the pullback of the Uhlmann connection by the family of single particle states parametrised by momenta is Abelian. The resulting connection is precisely the so-called Berry connection.

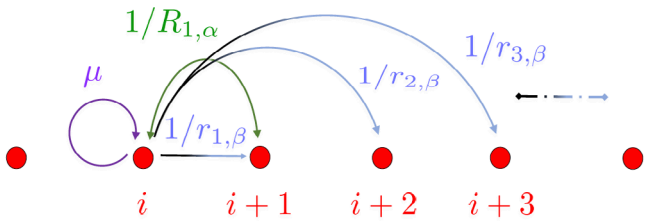


Figure 1. Depiction of 1D modified Kitaev superconductor Hamiltonian given in equation (11) with long-range hopping amplitudes (blue arrows), strictly short range pairing amplitudes (green arrow) and chemical potential (purple arrow).

We end this section with a remark on the quantity Δ which for the case of pure states can be algebraically computed in terms of the fidelity. More concretely, let $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\phi\rangle\langle\phi|$ be pure states, then

$$\Delta(\rho,\sigma) = F(\rho,\sigma) - \mathrm{Tr}\big(\sqrt{\rho}\sqrt{\sigma}\big) = F(\rho,\sigma) - F^2(\rho,\sigma). \tag{9}$$

For the same reasons as described for the fidelity, instead of the above pure state expression, we will use the zero temperature limit of (7).

3. Results

3.1. 1D topological superconductors with exponential decaying hopping amplitude

The Kitaev chain [12] is a one-dimensional model for topological superconductivity, where the hoppings and pairing amplitudes are between nearest neighbours. Various attempts have been made to modify the Kitaev chain by deforming the pairing and hopping amplitudes. We consider the following Hamiltonian with long range couplings [23]

$$\mathcal{H} = \sum_{i=1}^{N}\left(\frac{1}{2}\sum_{l=1}^{N-1}(-\frac{1}{r_{l,\beta}}c_{i+l}^\dagger c_i + \frac{1}{R_{l,\alpha}}c_{i+l}^\dagger c_i^\dagger + \mathrm{H.c.}) - \mu(c_i^\dagger c_i - \frac{1}{2})\right), \tag{10}$$

where $r_{l,\beta}$ and $R_{l,\alpha}$, control hopping and pairing amplitudes, respectively, and they are functions of the distance l, with real parameters β and α, respectively. The operators $c_i^\dagger$ (c_i) are spinless fermion creation (annihilation) operators and μ is the chemical potential. The short range pairing and long range hoppings in $\mathcal{H}$ are illustrated in figure 1. Observe that for strictly nearest-neighbour hopping and pairing amplitudes, i.e. for $1/r_{l,\beta} = 2t\delta_{l,1}$ and $1/R_{l,\alpha} = 2\Delta\delta_{l,1}$, we recover the original Kitaev chain Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{N}\left(-tc_{i+1}^\dagger c_i + \Delta c_{i+1}^\dagger c_i^\dagger + \mathrm{H.c.} - \mu(c_i^\dagger c_i - \frac{1}{2})\right). \tag{11}$$

Taking advantage of translational invariance, we can diagonalise the above Hamiltonian by going to Fourier space, obtaining a discrete version of equation (4),

$$\mathcal{H}(q) = \frac{1}{2N}\sum_k \widetilde{\psi}_k^\dagger H_{\mathrm{BdG}}(k;q)\widetilde{\psi}_k, \tag{12}$$

with $k \in \{\frac{2\pi\ell}{N} : \ell \in \{0,...,N-1\}\}$, $q = (\mu,\beta,\alpha)$, and

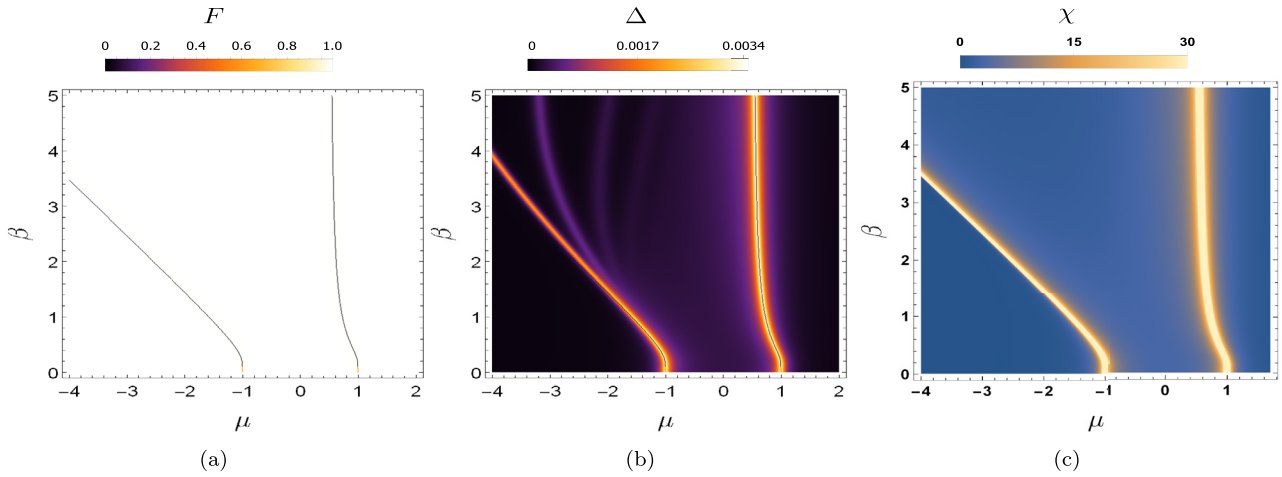


Figure 2. Fidelity $F(\rho,\rho')$ and (b) Uhlmann quantity $\Delta(\rho,\rho')$ for states $\rho=\rho(\mu,\beta)$ and $\rho'=\rho(\mu+\delta\mu,\beta)$ as functions of (μ,β), where $\delta\mu=0.01$ and $T=10^{-5}$ (c) Ground state fidelity susceptibility $\chi_{\mu\mu}=\chi_{\mu\mu}(\mu,\beta)$ for 1D TSCs as given by equation (11).

$$E(k;q)=\sqrt{(\mu+g_\beta(k))^2+f_\alpha^2(k)},$$
$$n(k;q)=-\frac{1}{E(k;q)}(0,f_\alpha(k),(\mu+g_\beta(k))),$$

where

$$f_\alpha(k)=\sum_{l=1}^{N-1}\frac{\sin(kl)}{R_{l,\alpha}}\qquad g_\beta(k)=\sum_{l=1}^{N-1}\frac{\cos(kl)}{r_{l,\beta}}.$$

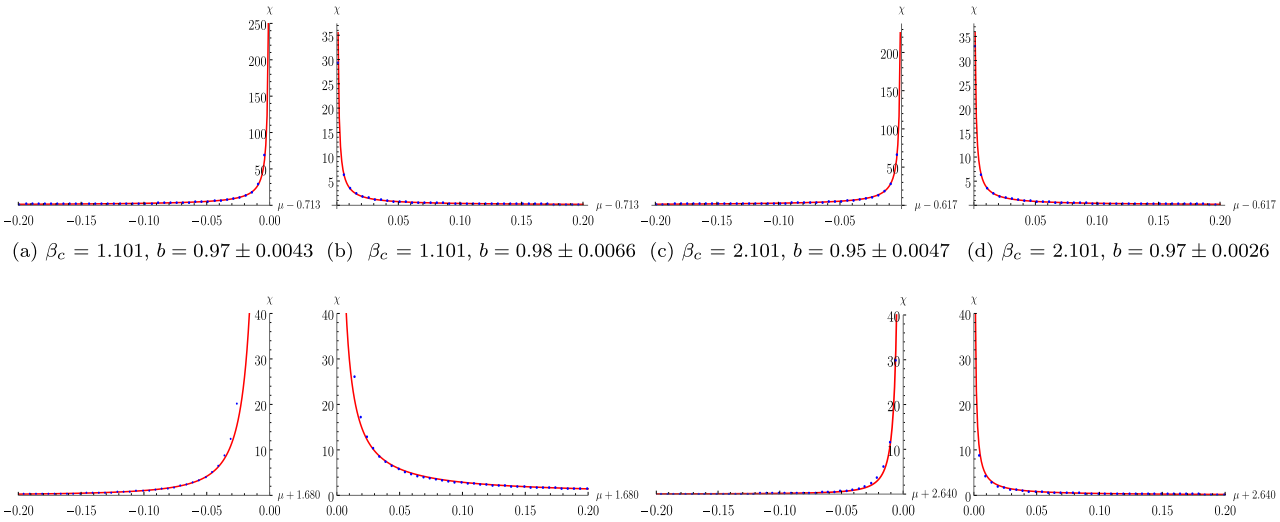
We study a special case where pairing is considered to be strictly nearest-neighbour, $1/R_{l,\alpha}=\delta_{1,l}$, and hopping terms which decay exponentially of the form $r_{l,\beta}=\exp\big((l-1)/\beta\big)$. This fixes the parameter α and hence we restrict ourselves to the parameters $q\equiv(\mu,\beta)$.

In the case of a 1D superconductor with strictly nearest hopping and pairing amplitudes, there is one nontrivial topological phase with winding number 1 for $-1<\mu<1$, where for other values of μ the winding number is 0. On the other hand, when long-range effects are included, i.e. $\beta>0$, the topological phase diagram, given in terms of μ, changes [23]. Our fidelity and Δ analysis show this exact behaviour, see figure 2. We probe the parameter space by considering two nearby points (μ,β) and $(\mu+\delta\mu,\beta)$ and calculate the fidelity $F=F(\rho(\mu,\beta),\rho(\mu+\delta\mu,\beta))$ and $\Delta=\Delta(\rho(\mu,\beta),\rho(\mu+\delta\mu,\beta))$ for $\delta\mu=0.01$ and temperature $T=10^{-5}$. We refer the reader to the appendices of [27, 29] for details of the analytical derivations of the expressions for the fidelity, Δ and the ground state fidelity susceptibility, also applicable for the 2D case studied below. Figure 2(a) shows the critical lines of topological phase transitions. The value of $\delta\mu$ is chosen such that the two states $\rho(\mu)$ and $\rho(\mu+\delta\mu)$ are generically the same. However, near the critical points, no matter how small $\delta\mu$ is, they are substantially different. Indeed, we can see that F is one everywhere except for the phase transition lines where its value drops down from one.

Unlike generic models where the fidelity between two different phases becomes zero only in the thermodynamic limit, in the current model, the zero temperature fidelity between two ground states from different phases is exactly zero, no matter how small the parameter difference is. This follows from the fact that at the special momentum points $k=0,\pi$, the vector $n(k)$ is proportional to the z axis and it happens that, for $k=\pi$, it has opposite signs in the different phases leading to a zero factor in the product determining the ground state fidelity. Indeed, when calculating the zero temperature pure state fidelity of a finite system we obtain the *exact* zero value in the vicinity of the critical point. Recall that at the critical points, the ground state manifold is degenerate; thus, we avoid such points from this calculation. To observe the exact zero value for fidelity we need to compare states from different phases. To ensure that this will always happen, the parameter difference, say $\delta\mu$, has to be bigger than the spacing between the two consecutive parameter points in which we calculate the ground state fidelity (grid spacing $\Delta\mu$); in our case, $\delta\mu=0.01>0.0099=\Delta\mu$. We would like to stress that this example shows the generality of the fidelity approach, which allowed us to correctly infer the critical points even without considering physical details of the system.

Figure 2(b) shows the behaviour of Δ: the departure of Δ from zero shows the nontrivial behaviour along the phase transition lines. Note that the faded curves near the left transition line represent numerical instabilities, as they do not feature in the plots of F and χ. The numerical value used for the number of lattice sites in the calculation of fidelity and the Uhlmann quantity is $N=80$. In figure 2(c) we plot the $\chi_{\mu\mu}$ which also shows the topological phase transitions. We can see from all the above plots that as β goes to zero we recover the usual topological superconductor where criticality exist at $\mu=1$ and -1, respectively. On the other hand, when β increases above zero we see the above mentioned drift of the critical points in the parameter space.

Along with the above analysis, we also performed asymptotic analysis of the ground state fidelity susceptibility in the neighbourhood (both left and right) of the critical points, where we obtain the associated critical exponents. The critical exponents are calculated as done in statistical mechanics and in the theory of phase transitions, where one wants to



(a) $\beta_c = 1.101$, $b = 0.97 \pm 0.0043$ (b) $\beta_c = 1.101$, $b = 0.98 \pm 0.0066$ (c) $\beta_c = 2.101$, $b = 0.95 \pm 0.0047$ (d) $\beta_c = 2.101$, $b = 0.97 \pm 0.0026$

(e) $\beta_c = 1.101$, $b = 1.98 \pm 0.0039$ (f) $\beta_c = 1.101$, $b = 1.01 \pm 0.0696$ (g) $\beta_c = 2.101$, $b = 1.90 \pm 0.0239$ (h) $\beta_c = 2.101$, $b = 1.02 \pm 0.0147$

Figure 3. The ground state fidelity susceptibility $\chi_{\mu\mu}$ as a function of $(\mu - \mu_c)$ with fixed values of $\beta = \beta_c = 1.101$ for (a), (b), (e), (f) and $\beta = \beta_c = 2.101$ for (c)–(f) in the neighbourhood of different critical μ_c: (a) and (b) $\mu_c = 0.713$, (c) and (d) $\mu_c = 0.617$, (e) and (f) $\mu_c = -1.68$, (g) and (h) $\mu_c = -2.64$. The points and red curve are the result of numerical integration and of the fit respectively.

understand the nature of the singular point (see for instance [36]). Near a gapless point, only the behaviour near the critical momenta is relevant. Therefore, there should exist some underlying universal behaviour captured in the correlation functions of theory. The fidelity susceptibility can be written in terms of the correlation functions (see section IV B of [37]) and therefore would also manifest this universality properties. The power law behaviour is typical for gapless systems, due to generic arguments of conformal invariance.

A least squares fit of the susceptibility to a power law $\chi \propto |\mu - \mu_c|^{-b}$ in a neighbourhood of μ_c was performed for different critical β_c, as shown in figure 3. In obtaining the parameter b and the associated uncertainty, we used the standard formulas from linear regression (see, for instance, [38], p 462) on a log–log plot. For $\beta_c = 1.101$, and we found:

(i) in the neighborhood of $\mu_c = 0.713$, exponents $b = 0.97 \pm 0.0043$ from the left, see figure 3(a), and $b = 0.98 \pm 0.0066$ from the right, see figure 3(b).
(ii) in the neighborhood of $\mu_c = -1.680$, exponents $b = 1.98 \pm 0.0039$ from the left, see figure 3(e) and $b = 1.01 \pm 0.0696$ from the right, see figure 3(f).

Similarly, for $\beta_c = 2.101$, we found:

(i) in the neighbourhood of $\mu_c = 0.617$, exponents $b = 0.95 \pm 0.0047$ from the left, see figure 3(c), and $b = 0.97 \pm 0.0026$ from the right, see figure 3(d).
(ii) in the neighbourhood of $\mu_c = -2.640$, exponents $b = 1.90 \pm 0.0339$ from the left, see figure 3(g), and $b = 1.02 \pm 0.0147$ from the right, see figure 3(h).

3.2. 2D topological superconductors with long-range effects

In the same way as in case of 1D TSCs, modifications are made in the hopping and pairing amplitudes of short range 2D chiral superconductors. We consider the following modified Hamiltonian model [24]

$$\mathcal{H} = \sum_{m,l} \left(\sum_{r,s}{}' \left(-\frac{t}{d_{r,s}^{\beta}} c_{l+r,m+s}^{\dagger} c_{l,m} + \frac{\Delta}{d_{r,s}^{\alpha+1}} (r + \mathrm{i}s) \right.\right.$$
$$\left.\left. c_{l+r,m+s}^{\dagger} c_{l,m}^{\dagger}\right) - (\mu - 4t) c_{l,m}^{\dagger} c_{l,m} \right) + \text{H.c.}, \tag{13}$$

where $\sum'_{r,s}$ means that $r, s = 0$ are excluded from the sum. Again the $c_i^{\dagger}$'s (c_i's) are spinless fermion creation (annihilation) operators. The hopping amplitude decays as $1/d_{r,s}^{\beta}$ and the pairing amplitude as $1/d_{r,s}^{\alpha}$, where $d_{r,s} = \sqrt{r^2 + s^2}$ is the Euclidean distance between the coupled sites. One can recover the usual short range 2D chiral p-wave superconductor Hamiltonian from above model by restricting amplitudes to only nearest neighbours. Observe that the r and s indices control the hoppings and pairings along the x and y directions, respectively. If we restrict hopping and pairing amplitudes to nearest neighbours only, this would be equivalent to consider only $(r, s) \in \{(1, 0), (0, 1)\}$. This yields

$$\mathcal{H} = \sum_{l,m} \left(\left(-\frac{t}{d_{1,0}^{\beta}} c_{l+1,m}^{\dagger} c_{l,m} - \frac{t}{d_{0,1}^{\beta}} c_{l,m+1}^{\dagger} c_{l,m} + \frac{\Delta}{d_{1,0}^{\alpha+1}} c_{l+1,m}^{\dagger} c_{l,m}^{\dagger} \right.\right.$$
$$\left.\left. + \mathrm{i}\frac{\Delta}{d_{0,1}^{\alpha+1}} c_{l,m+1}^{\dagger} c_{l,m}^{\dagger}\right) - (\mu - 4t) c_{l,m}^{\dagger} c_{l,m} \right) + \text{H.c.} \tag{14}$$

Setting $d_{0,1} = d_{1,0} = 1$, we recover the well known short range chiral p-wave superconductors [39–42] described by the following Hamiltonian

$$\mathcal{H} = \sum_{l,m} \left(-t(c_{l+1,m}^{\dagger} c_{l,m} + c_{l,m+1}^{\dagger} c_{l,m}) \right. \tag{15}$$
$$\left. + \Delta(c_{l+1,m}^{\dagger} c_{l,m}^{\dagger} + \mathrm{i} c_{l,m+1}^{\dagger} c_{l,m}^{\dagger}) - (\mu - 4t) c_{l,m}^{\dagger} c_{l,m} \right) + \text{H.c.}$$

Note that the same Hamiltonian can be obtained by taking α and β to $+\infty$. In this limit only the terms $d_{1,0} = d_{0,1} = 1$

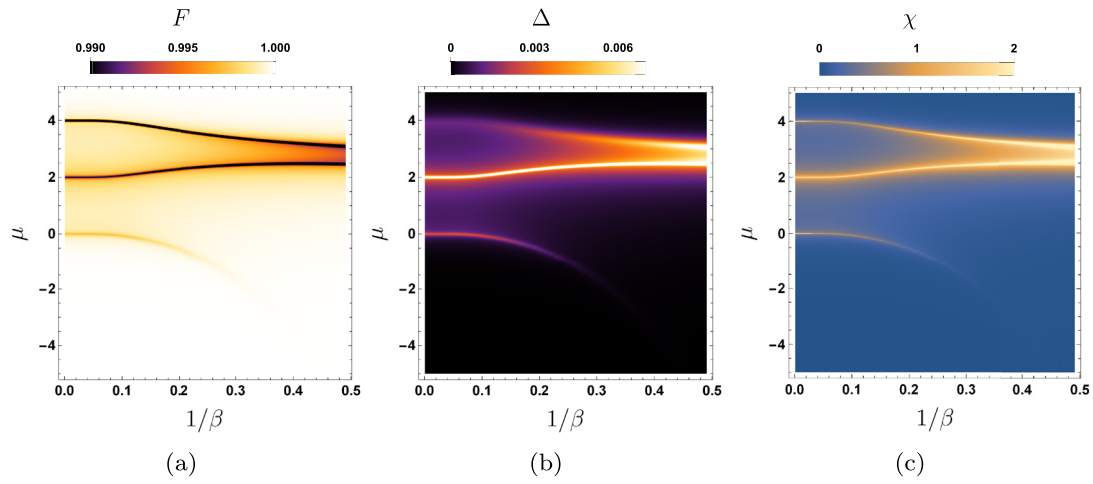
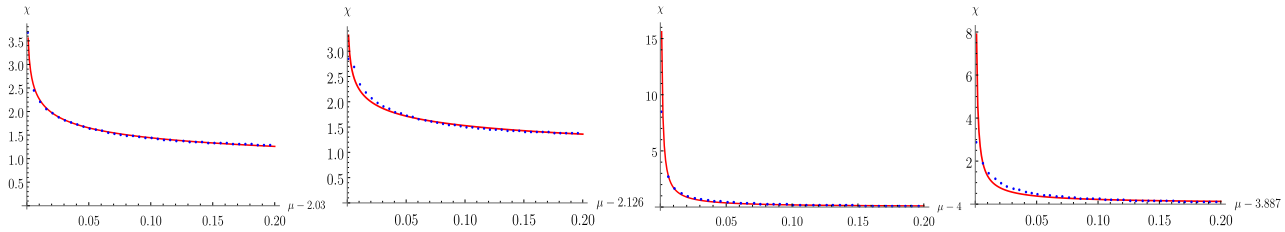


Figure 4. (a) Fidelity $F(\rho,\rho')$ and (b) Uhlmann quantity $\Delta(\rho,\rho')$ for states $\rho=\rho(\mu,\beta)$ and $\rho'=\rho(\mu+\delta\mu,\beta)$ as functions of (μ,β), where $\delta\mu=0.01$ and $T=6\times10^{-3}$. Note that, just like in the 1D case, if we were to choose smaller temperatures and finer grid spacings, we would have observed more prominent behaviour of F and Δ along the lines of phase transitions. Due to higher complexity of the 2D problem and the rather limited computational power available to us, we omitted such finer analysis. (c) Ground state fidelity susceptibility $\chi_{\mu\mu}=\chi_{\mu\mu}(\mu,\beta)$ for 2D TSCs as given by equation (14).



(a) $1/\beta_c=0.0181$, $b=0.19\pm0.0014$ (b) $1/\beta_c=0.1223$, $b=0.17\pm0.0050$ (c) $1/\beta_c=0.0181$ $b=0.96\pm0.0236$ (d) $1/\beta_c=0.1223$, $b=0.79\pm0.0370$

Figure 5. The ground state fidelity susceptibility $\chi_{\mu\mu}$ as a function of $\mu-\mu_c$ with fixed values of $1/\beta=1/\beta_c=0.0181$ for (a), (c) and $1/\beta=1/\beta_c=0.1223$ for (b) and (d) in the neighbourhood of different critical μ_c:(a) $\mu_c=2.03$, (b) $\mu_c=2.126$, (c) $\mu_c=4$ and (d) $\mu_c=3.887$.

from the primed sum survive, and the usual short-range superconductors are obtained [24]. Here onward we take $\Delta=t=1/2$. Considering periodic boundary conditions, i.e. putting our system on a torus, the Hamiltonian in equation (14) can be brought into the form given in equation (4) with

$$H_{\mathrm{BdG}}(\mathbf{k})=f_\alpha(\mathbf{k})\sigma_x+h_\alpha(\mathbf{k})\sigma_y-2\big[\mu-4t+tg_\beta(\mathbf{k})\big]\sigma_z, \tag{16}$$

where

$$f_\alpha(\mathbf{k})=\sum_{r,s}{}'\frac{s\sin(k_xr+k_ys)}{d_{r,s}^{\alpha+1}},\ h_\alpha(\mathbf{k})=\sum_{r,s}{}'\frac{r\sin(k_xr+k_ys)}{d_{r,s}^{\alpha+1}},$$
$$g_\beta(\mathbf{k})=\sum_{r,s}{}'\frac{\cos(k_xr+k_ys)}{d_{r,s}^{\beta}}, \tag{17}$$

and $\mathbf{k}=(k_x,k_y)$. It is known [24] that in the case of the short-range 2D TSC described above, in the limit $\alpha,\beta\to+\infty$, there exist two non-trivial TPs with Chern numbers $C=1$, for $0<\mu<2$, and $C=-1$, for $2<\mu<4$. The points of quantum phase transition are given by $\mu_c=0,2,4$. Switching on the long-range effects, the critical points start drifting, and as a result the nontrivial topological phase with $C=-1$ gets suppressed, while the $C=1$ phase gets enlarged.

The plots for the fidelity, Δ and the ground state fidelity susceptibility $\chi_{\mu\mu}$ are shown in figure 4, where we take $\beta=\alpha$. The fidelity and Δ are obtained between two close states $\rho(\mu,\beta)$ and $\rho(\mu+\delta\mu,\beta)$ for $\delta\mu=0.01$, temperature $T=6\times10^{-3}$ and grid spacing $\Delta\mu=0.02$. In numerical calculations of the fidelity and Δ, the number of sites along the x (y) direction is taken to be $N_x=16$ ($N_y=16$). The departure of fidelity from 1 and Δ from 0 signals the quantum phase transition lines, as shown in figures 4(a) and (b), respectively. The ground state fidelity susceptibility also shows the phase transition lines in the parameter space, see figure 4(c). When calculating the zero temperature pure state fidelity of a finite system we obtain, as in the 1D case, the exact zero value in the vicinity of the critical point. The reason is because the Bogoliubov eigenmodes in different phases are not adiabatically connected. Therefore there exists at least one momentum for which the overlap between these modes is zero.

We also performed the asymptotic analysis of the ground state fidelity susceptibility in the neighbourhood of critical μ_c for fixed value of $\beta=\beta_c$ as shown in figure 5. For

$1/\beta_c = 0.0181$, we performed a least squares fit of the susceptibility to a power law $\chi \propto |\mu - \mu_c|^{-b}$, yielding:

(i) in the neighbourhood of $\mu_c = 2.03$, exponent $b = 0.19 \pm 0.0014$ from the right, see figure 5(a).
(ii) in the neighbourhood of $\mu_c = 4.00$, exponent $b = 0.96 \pm 0.0181$ from the right, see figure 5(c).

Similarly, for $1/\beta_c = 0.1223$, the least squares fit yields:

(i) in the neighbourhood of $\mu_c = 2.13$, exponent $b = 0.17 \pm 0.0050$ from the right, see figure 5(b).
(ii) in the neighbourhood of $\mu_c = 3.89$, exponent $b = 0.79 \pm 0.0370$ from the right, see figure 5(d).

4. Conclusions

In this work we performed an information geometric analysis of 1D and 2D topological superconductors with long-range hopping and pairing amplitudes. The fidelity, Δ and the ground state fidelity susceptibility confirmed the phase diagram as obtained in the recent works [23, 24]. We also performed the analysis of the critical exponents for the ground state fidelity susceptibility. Our analysis in terms of fidelity allows for a natural extension of the study to finite temperatures, where the notion of topological phase is yet to be completely understood. Moreover, it would be interesting to understand the interplay between long range phenomena and temperature fluctuations. Traditionally, the impact of thermal fluctuations is often studied in terms of local order parameters. In the case of topological phase transitions, the role of order parameters could be assigned to global topological invariants, but they are applicable only at the zero temperature. Thus, finding possible observable quantities that could probe topological orders at finite temperatures is an important open problem. The fidelity approach offers one such candidate: an *optimal observable* for the quantum fidelity. In short, measuring an optimal observable on the systems in states ρ and σ produces two classical probability distributions whose classical fidelity is equal to the quantum fidelity between ρ and σ (for details, see [43], section 9.2.2, p 412). We believe analysing such possibility could possibly be an intriguing line of future research.

Acknowledgments

STA, BM and NP acknowledge the support of SQIG—Security and Quantum Information Group, under the Fundação para a Ciência e a Tecnologia (FCT) project UID/EEA/50008/2019, and European funds, namely H2020 project SPARTA. STA acknowledges the support from DP-PMI and FCT (Portugal) through the Grant PD/ BD/113651/2015. BM and NP acknowledge projects QuantMining POCI-01-0145-FEDER-031826, PREDICT PTDC/CCI-CIF/29877/2017 and internal IT project, QBigData PEst-OE/EEI/LA0008/2013, funded by FCT. NP thanks FCT Grant CEECIND/04594/2017. Support from FCT (Portugal) through Grant UID/CTM/04540/2013 is also acknowledged.

ORCID iDs

Bruno Mera https://orcid.org/0000-0002-7214-6506

References

[1] Thouless D J, Kohmoto M, Nightingale M P and den Nijs M 1982 Quantized Hall conductance in a two-dimensional periodic potential *Phys. Rev. Lett.* **49** 405–8
[2] Osterloh G F A, Amico L and Fazio R 2002 Scaling of entanglement close to a quantum phase transition *Nature* **416** 608–10
[3] Su S-Q, Song J-L and Gu S-J 2006 Local entanglement and quantum phase transition in a one-dimensional transverse field Ising model *Phys. Rev.* A **74** 032308
[4] Vidal G, Latorre J I, Rico E and Kitaev A 2003 Entanglement in quantum critical phenomena *Phys. Rev. Lett.* **90** 227902
[5] Oliveira T P and Sacramento P D 2014 Entanglement modes and topological phase transitions in superconductors *Phys. Rev.* B **89** 094512
[6] Uhlmann A 1976 The 'transition probability' in the state space of a C*-algebras *Rep. Math. Phys.* **9** 273–9
[7] Alberti P M 1983 A note on the transition probability over C*-algebras *Lett. Math. Phys.* **7** 25–32
[8] Quan H T, Song Z, Liu X F, Zanardi P and Sun C P 2006 Decay of Loschmidt echo enhanced by quantum criticality *Phys. Rev. Lett.* **96** 140604
[9] Zanardi P and Paunković N 2006 Ground state overlap and quantum phase transitions *Phys. Rev.* E **74** 031123
[10] Paunković N, Sacramento P D, Nogueira P, Vieira V R and Dugaev V K 2008 Fidelity between partial states as a signature of quantum phase transitions *Phys. Rev.* A **77** 052302
[11] Sacramento P D, Paunković N and Vieira V R 2011 Fidelity spectrum and phase transitions of quantum systems *Phys. Rev.* A **84** 062318
[12] Kitaev A Y 2001 Unpaired majorana fermions in quantum wires *Phys.—Usp.* **44** 131–6
[13] Patrick K, Neupert T and Pachos J K 2017 Topological quantum liquids with long-range couplings *Phys. Rev. Lett.* **118** 267002
[14] Pientka F, Glazman L I and von Oppen F 2013 Topological superconducting phase in helical shiba chains *Phys. Rev.* B **88** 155420
[15] Klinovaja J, Stano P, Yazdani A and Loss D 2013 Topological superconductivity and majorana fermions in RKKY systems *Phys. Rev. Lett.* **111** 186805
[16] Pientka F, Glazman L I and von Oppen F 2014 Unconventional topological phase transitions in helical shiba chains *Phys. Rev.* B **89** 180505
[17] Neupert T, Yazdani A and Bernevig B A 2016 Shiba chains of scalar impurities on unconventional superconductors *Phys. Rev.* B **93** 094508
[18] Niu Y, Chung S B, Hsu C-H, Mandal I, Raghu S and Chakravarty S 2012 Majorana zero modes in a quantum ising chain with longer-ranged interactions *Phys. Rev.* B **85** 035110
[19] Vodola D, Lepori L, Ercolessi E, Gorshkov A V and Pupillo G 2014 Kitaev chains with long-range pairing *Phys. Rev. Lett.* **113** 156402
[20] Dutta A and Dutta A 2017 Probing the role of long-range interactions in the dynamics of a long-range kitaev chain *Phys. Rev.* B **96** 125113
[21] Alecce A and Dell'Anna L 2017 Extended Kitaev chain with longer-range hopping and pairing *Phys. Rev.* B **95** 195160
[22] Lepori L and Dell'Anna L 2017 Long-range topological insulators and weakened bulk-boundary correspondence *New J. Phys.* **19** 103030

[23] Viyuela O, Vodola D, Pupillo G and Martin-Delgado M A 2016 Topological massive Dirac edge modes and long-range superconducting hamiltonians *Phys. Rev.* B **94** 125121
[24] Viyuela O, Fu L and Martin-Delgado M A 2018 Chiral topological superconductors enhanced by long-range interactions *Phys. Rev. Lett.* **120** 017001
[25] Ruby M, Heinrich B W, Peng Y, von Oppen F and Franke K J 2017 Exploring a proximity-coupled co chain on pb(1 1 0) as a possible majorana platform *Nano Lett.* **17** 4473–7
[26] de Bruijckere J, Ugeda M M, Lorente N, Pascual J I, Choi D-J and Rubio-Verdú C 2017 Mapping the orbital structure of impurity bound states in a superconductor *Nat. Commun.* **8** 15175
[27] Mera B, Vlachou C, Paunković N and Vieira V R 2017 Uhlmann connection in fermionic systems undergoing phase transitions *Phys. Rev. Lett.* **119** 015702
[28] Mera B, Vlachou C, Paunković N and Vieira V R 2017 Boltzmann–Gibbs states in topological quantum walks and associated many-body systems: fidelity and Uhlmann parallel transport analysis of phase transitions *J. Phys. A: Math. Theor.* **50** 365302
[29] Amin S T, Mera B, Vlachou C, Paunković N and Vieira V R 2018 Fidelity and Uhlmann connection analysis of topological phase transitions in two dimensions *Phys. Rev.* B **98** 245141
[30] Paunković N and Rocha Vieira V 2008 Macroscopic distinguishability between quantum states defining different phases of matter: fidelity and the Uhlmann geometric phase *Phys. Rev.* E **77** 011129
[31] Anderson P W 1967 Infrared catastrophe in Fermi gases with local scattering potentials *Phys. Rev. Lett.* **18** 1049
[32] Gu J and Sun K 2016 Adiabatic continuity, wave-function overlap, and topological phase transitions *Phys. Rev.* B **94** 125111
[33] Huang Z and Balatsky A V 2016 Dynamical quantum phase transitions: role of topological nodes in wave function overlaps *Phys. Rev. Lett.* **117** 086802
[34] Chruscinski D and Jamiolkowski A 2012 *Geometric Phases in Classical and Quantum Mechanics* vol 36 (Berlin: Springer) (https://doi.org/10.1007/978-0-8176-8176-0)
[35] Uhlmann A 1989 On Berry phases along mixtures of states *Ann. Phys., Lpz.* **501** 63–9
[36] Anderson P W 1997 *Basic Notions of Condensed Matter Physics* (Boulder, CO: Westview Press)
[37] Gu S-J 2010 Fidelity approach to quantum phase transitions *Int. J. Mod. Phys.* B **24** 4371–458
[38] Lane D M, Scott D, Hebl M, Guerra R, Osherson D and Zimmer H 2017 *Introduction to Statistics* (Houston, TX: Rice University)
[39] Bernevig B A and Hughes T L 2013 *Topological Insulators and Topological Superconductors* (Princeton, NJ: Princeton University Press) (https://doi.org/10.1515/9781400846733)
[40] Ivanov D A 2001 Non-abelian statistics of half-quantum vortices in *p*-wave superconductors *Phys. Rev. Lett.* **86** 268–71
[41] Read N and Green D 2000 Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum hall effect *Phys. Rev.* B **61** 10267–97
[42] Volovik G E 1999 Fermion zero modes on vortices in chiral superconductors *J. Exp. Theor. Phys. Lett.* **70** 609–17
[43] Nielsen M A and Chuang I L 2002 *Quantum Information and Quantum Computation* (Cambridge: Cambridge University Press) (https://doi.org/10.1017/CBO9780511976667)

International Journal of Modern Physics C
Vol. 28, No. 4 (2017) 1750055 (27 pages)
© World Scientific Publishing Company
DOI: 10.1142/S0129183117500553

World Scientific
www.worldscientific.com

A simulator for discrete quantum walks on lattices

J. Rodrigues*, N. Paunković† and P. Mateus‡

Instituto de Telecomunicações,
DM, Instituto Superior Técnico, University of Lisbon,
Av. Rovisco Pais 1049–001, Lisboa, Portugal
**ler@technico.ulishoa.pt*
†*npaunkov@math.ist.utl.pt*
‡*paulo.mateus@lx.it.pt*

Received 13 April 2015
Revised 22 February 2016
Accepted 9 February 2017
Published 11 April 2017

In this paper, we present a simulator for two-particle quantum walks on the line and one-particle on a two-dimensional squared lattice. It can be used to investigate the equivalence between the two cases (one- and two-particle walks) for various boundary conditions (open, circular, reflecting, absorbing and their combinations). For the case of a single walker on a two-dimensional lattice, the simulator can also implement the Möbius strip. Furthermore, other topologies for the walker are also simulated by the proposed tool, like certain types of planar graphs with degree up to 4, by considering missing links over the lattice. The main purpose of the simulator is to study the genuinely quantum effects on the global properties of the two-particle joint probability distribution on the entanglement between the walkers/axis. For that purpose, the simulator is designed to compute various quantities such as: the entanglement and classical correlations, (classical and quantum) mutual information, the average distance between the two walkers, different hitting times and quantum discord. These quantities are of vital importance in designing possible algorithmic applications of quantum walks, namely in search, 3-SAT problems, etc. The simulator can also implement the static partial measurements of particle(s) positions and dynamic breaking of the links between certain nodes, both of which can be used to investigate the effects of decoherence on the walker(s). Finally, the simulator can be used to investigate the dynamic Anderson-like particle localization by varying the coin operators of certain nodes on the line/lattice. We also present some illustrative and relevant examples of one- and two-particle quantum walks in various scenarios. The tool was implemented in C and is available on-line at http://qwsim.weebly.com/.

Keywords: Quantum walks; simulator; broken links; boundary conditions; entanglement.

PACS Nos.: 03.67–a, 03.67.Ac

1. Introduction

Quantum walks[1] are the quantum analog of classical random walks. They have been used both to develop quantum-based algorithms and to model various physical

‡Corresponding author.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

systems and processes. It was shown that under certain assumptions and scenarios, continuous-time quantum walks can be exponentially faster than classical ones, see Refs. 2 and 3. They have been used in designing search algorithms,[4] and showed to be quadratically faster than any known classical counterparts.[5] Discrete-time quantum walks have also been shown to be useful in searching spatial regions under realistic assumptions, and moreover, to clearly surpass, at least quadratically, common search algorithms, see Refs. 6–11. Similar results were obtained when quantizing classical algorithms based on random walks, namely when the underlying Markov chain is ergodic and symmetric,[10] or even to a larger class of Markov chains.[11]

Multi-particle quantum walks were first introduced in Ref. 12 where it was shown that certain joint two-particle properties, such as the mean distance between the walkers and the hitting time for *at least one* walker, can be for walkers initially entangled between each other larger or smaller than in the case of unentangled walkers. Further work studied the meeting problem[13] and the generation of the entanglement between the walkers,[14] confirming the genuine quantum mechanical influence of entanglement on the joint features of quantum walks, which in turn might affect the performance of the above-mentioned computational applications.

Analytical results concerning quantum walks, like hitting, mixing and dispersion time (among others), were extensively studied for single-particle quantum walks (see for example Refs. 15 and 16). Also, the tools for the numerical study of one-particle quantum walks in one and two dimensions have been presented before (see Refs. 17 and 18). Yet, although the entanglement between the coin and the position degrees of freedom have been studied in detail both numerically and analytically (see for example Refs. 19 and 20, respectively), a very little work has been done concerning the effects of the entanglement between the walkers. In order to address these issues, we designed a tool for numerical simulation of two-particle quantum walks.

Herein, we present the main result of the paper — a simulator for two-particle quantum walks in one dimension and one-particle quantum walks in two dimensions. The simulator computes quantities that measure the global features that are influenced by the entanglement between the two walkers, such as correlations, classical (Shannon) and quantum (von Neumann) mutual information between the two walkers, average distance between the walkers, various hitting times for *at least one* walker, etc. It can include various boundary conditions — open, reflecting and absorbing — and combine them for each particle independently, resulting in different two-particle geometries, such as torus, cylinder, finite squared lattice with reflecting boundaries, etc. For one particle on a two-dimensional lattice, a new boundary conditions resulting in a Möbious strip and Klein bottle topologies are possible to consider, which would in the case of two particles in one dimension require non-local shift operation for at least one particle. By implementing the static partial position measurements (constant during the walk) and/or dynamic breaking links between certain nodes (random choice of links broken at each step of the walk), our simulator can be used to study the effects of decoherence. The dynamic change of coin operators can also be implemented and used to study various types of particle

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

localizations,[22] such as those of Anderson type and others.[22] The simulator can perform up to 10^7 steps of a discrete-time quantum walk on the open line. The memory needed to encode an arbitrary state of such walk after 10^8 steps would be around 12 000 Gbits of RAM, which exceeds the memory of standard computers available in the market. Thus our simulator saturates the current technological limits of an average consumer. When restricting the memory needed to encode the walk by considering reflecting boundaries in the nodes, say -4000 and $+4000$, the simulator performs 10^8 steps in about two weeks on a standard 3 GHz computer, which is reasonable for a well-defined scientific research.

Other results, presented as examples in Sec. 4, confirm the expected dependence of the above-mentioned joint quantities on entanglement, in accordance with Theorems 1–3 from Sec. 3. Also, we explicitly showed the examples of Anderson-like localizations in the presence of various static and dynamic decoherence mechanisms (i.e. partial position measurements, broken links and varying coin operations).

The paper is organized as follows. After this introduction, in Sec. 2 we present a mathematical description of a discrete-time quantum walk on a line with one and two particles, and show the equivalence of the latter with a one-particle quantum walk on a square lattice. We discuss different boundary conditions (circular, reflecting, absorbing, etc.) and for the case of a one-particle walk on a lattice, different topologies (Möbius strip, Klein bottle, etc.). Finally, we describe two models for a noise in a quantum walk: dynamic breaking of links between certain nodes, and varying coins for random nodes. In Sec. 3, we describe the quantities that our simulator is calculating and analyze their relevance, with the emphasis to a two-particle quantum walk picture where joint properties depending on correlation and entanglement can exhibit specific nonclassical quantum features. In Sec. 4, we present some illustrative examples. Finally, in Sec. 5, we present conclusions and possible future improvements to the code.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

2. Discrete-Time Quantum Walk

In a discrete-time quantum walk on a line, we consider the movement of a walker along discrete positions, labeled on a line $x \in \mathbb{Z}$. At each step this particle can move to the left or to the right of the line. The direction is controlled by an internal degree of freedom, commonly called the coin degree of freedom. Both position and coin states of a given particle can be modeled using Hilbert spaces $\mathcal{H}_P = \text{span}\{|x\rangle : x \in \mathbb{Z}\}$ and $\mathcal{H}_C = \text{span}\{|R\rangle, |L\rangle\}$, for the position space and the coin space, respectively. The total Hilbert space of a particle doing a discrete-time quantum walk on a line is given by $\mathcal{H} = \mathcal{H}_P \otimes \mathcal{H}_C$. The one-step time evolution of the system is described by the unitary operator

$$\hat{U} = \hat{S}(\hat{I}_P \otimes \hat{U}_C), \tag{1}$$

where $\hat{S}$ is the shift operator given by

$$\hat{S} = \left(\sum_x |x+1\rangle\langle x|\right) \otimes |R\rangle\langle R| + \left(\sum_x |x-1\rangle\langle x|\right) \otimes |L\rangle\langle L|, \tag{2}$$

$\hat{I}_p$ is the identity operator on $\mathcal{H}_P$, and $\hat{U}_C \in \mathrm{U}(2)$ acts on $\mathcal{H}_C$.

Now, consider two noninteracting particles on a line. The joint Hilbert space of the composite system, consisting of two distinguishable particles 1 and 2 doing a quantum walk over the *same*[a] line, is

$$\mathcal{H}_{12} \equiv \mathcal{H}_1 \otimes \mathcal{H}_2,$$

where $\mathcal{H}_1 = \mathcal{H}_{P,1} \otimes \mathcal{H}_{C,1}$ and $\mathcal{H}_2 = \mathcal{H}_{P,2} \otimes \mathcal{H}_{C,2}$ represent the Hilbert spaces of particles 1 and 2, respectively. The joint one-step time evolution of this system is simply the tensor product between the unitary operators for time evolutions of each particle

$$\begin{aligned}\hat{U}_{12} = \hat{U}_1 \otimes \hat{U}_2 &= [\hat{S}_1(\hat{I}_{P,1} \otimes \hat{U}_{C,1})] \otimes [\hat{S}_2(\hat{I}_{P,2} \otimes \hat{U}_{C,2})] \\ &= \hat{S}_{12}([\hat{I}_{P,1} \otimes \hat{I}_{P,2}] \otimes \hat{U}_{C,12}),\end{aligned}$$

where $\hat{S}_{12} = \hat{S}_1 \otimes \hat{S}_2$ has the form

$$\begin{aligned}\hat{S}_{12} = \sum_{x_1,x_2} &|x_1+1, x_2+1\rangle_{12}\langle x_1, x_2| \otimes |RR\rangle_{12}\langle RR| \\ &+ |x_1+1, x_2-1\rangle_{12}\langle x_1, x_2| \otimes |RL\rangle_{12}\langle RL| \\ &+ |x_1-1, x_2+1\rangle_{12}\langle x_1, x_2| \otimes |LR\rangle_{12}\langle LR| \\ &+ |x_1-1, x_2-1\rangle_{12}\langle x_1, x_2| \otimes |LL\rangle_{12}\langle LL|\end{aligned} \tag{3}$$

and the joint coin operator is

$$\hat{U}_{C,12} = \hat{U}_{C,1} \otimes \hat{U}_{C,2}.$$

Note that the labels of ket states denote the joint and single-particle Hilbert spaces $\mathcal{H}_{12}$, $\mathcal{H}_1$ and $\mathcal{H}_2$, such that $|x_1, x_2\rangle_{12} \equiv |x_1\rangle_1 |x_2\rangle_2$, etc.

It is easy to see that a quantum walk of two particles on a line, in which initial positions of both walkers are equal, say 0, is equivalent to a quantum walk of one particle along a two-dimensional square lattice $x\mathrm{O}y$, whose nodes are labeled by their coordinates $\{(x, y) : x, y \in \mathbb{Z}\}$ along two perpendicular axes x and y. Indeed,

[a] From mathematical point of view, it is irrelevant whether two particles are performing the walk over the same or different lines: both descriptions are identical. Yet, it is crucial for the applications. For example, in the study of the effects of entanglement between the two walkers on the features of quantum walk-based search algorithms, in which the two walkers are performing a search over the *same* data-base. For the same reason, we assume that the walks are performed *simultaneously*, which is the reason for requiring particle distinguishability: otherwise, the states of two identical particles would be subject to bosonic and fermionic symmetrization and anti-symmetrization rules. On the other hand, such states are possible to study in the case of distinguishable particles as well.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

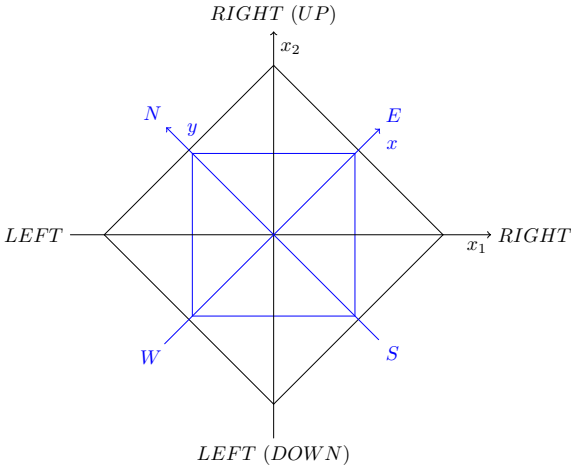


Fig. 1. Relation between the representations of quantum walks of one particle on a lattice and two particles on a line.

if the positions of the two particles on a line represent the two orthogonal coordinates, along the axes 1 and 2, of a node in the square lattice, then the two-particle configuration $(1,1)$ corresponds to a position $(1,0)$ of a particle on the $x\mathrm{O}y$ lattice whose axes x and y are rotated by $\pi/4$ with respect to the axes 1 and 2, see Fig. 1.

In general, the correspondence between a two-particle configuration (x_1, x_2) and a position (x, y) of a node in a rotated $x\mathrm{O}y$ lattice is given by

$$\begin{aligned} x &= \frac{1}{2}(x_2 + x_1), \\ y &= \frac{1}{2}(x_2 - x_1), \end{aligned} \tag{4}$$

which establishes the correspondence between the states from a two-particle position Hilbert space $(\mathcal{H}_{P,12} \equiv \mathcal{H}_{P,1} \otimes \mathcal{H}_{P,2})$ and a position Hilbert space $\mathcal{H}_{P,xy} = \mathrm{span}\{|x, y\rangle : x, y \in \mathbb{Z}\}$ of a single particle on a square lattice $x\mathrm{O}y$. According to this, moving both particles to the right along a line is equivalent to a particle on a lattice moving East (to the right with respect to the x-axis). This induces the following correspondence between the states from a two-particle coin Hilbert space $\mathcal{H}_{C,12} \equiv \mathcal{H}_{C,1} \otimes \mathcal{H}_{C,2}$ and the one-particle coin space $\mathcal{H}_{C,xy} = \mathrm{span}\{|E\rangle, |S\rangle, |N\rangle, |W\rangle\}$:

$$\begin{aligned} |E\rangle &= |RR\rangle_{12}, \quad |S\rangle = |RL_{12}\rangle, \\ |N\rangle &= |LR_{12}\rangle, \quad |W\rangle = |LL_{12}\rangle. \end{aligned} \tag{5}$$

Therefore, the overall Hilbert space of a single particle doing a quantum walk along the $x\mathrm{O}y$ squared lattice is $\mathcal{H}_{xy} \equiv \mathcal{H}_{P,xy} \otimes \mathcal{H}_{C,xy}$ (note that, for reasons of simplicity, we drop the labels of the ket states form $\mathcal{H}_{xy}$). The above correspondences (4) and (5)

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

give the shift operator equivalent to $\hat{S}_{12}$

$$\begin{aligned}\hat{S}_{xy} = \sum_{xy} & |x+1,y\rangle\langle x,y| \otimes |E\rangle\langle E| \\ & + |x,y-1\rangle\langle x,y| \otimes |S\rangle\langle S| \\ & + |x,y+1\rangle\langle x,y| \otimes |N\rangle\langle N| \\ & + |x-1,y\rangle\langle x,y| \otimes |W\rangle\langle W|, \end{aligned} \tag{6}$$

while the coin operator for a quantum walk of a particle on the lattice, equivalent to a *given* two-particle walk on a line, is unchanged: $\hat{U}_{C,xy} = \hat{U}_{C,12}$. Note that in general $\hat{U}_{C,xy} \in \mathrm{U}(4) \supset \mathrm{U}(2) \otimes \mathrm{U}(2)$, which in the case of two particles on a line would correspond to a global coin operation that could increase the entanglement between the coins of two particles. The unitary time evolution of each step of the quantum walk is given by

$$\hat{U}_{xy} = \hat{S}_{xy} \otimes (\hat{I}_{P,xy} \otimes \hat{U}_{C,xy}),$$

where $\hat{I}_{P,xy}$ is the identity operator on $\mathcal{H}_{P,xy}$.

Note that the above correspondence between a two-particle walk on a line and a single-particle walk on a square lattice is valid as long as initial positions of two particles on a line are of the same parity. Otherwise, using (4) we see that the two-particle walk is equivalent to a walk on a *different* square lattice whose node positions are *half-integers*. This way, we can use our simulator for two particles on a line to simulate quantum walks of two distinguishable particles on the *same* lattice, as long as their joint state is a product between one-particle states.

2.1. *Quantum walks with specific boundary conditions and topologies*

Quantum walks with certain boundary conditions are quite interesting. Firstly, from the theoretical point, it is essential to understand how these boundaries affect the features of quantum walks, in particular the entanglement between the particles as well as the other mentioned quantities that we study. Moreover, from a practical point of view, these boundaries might reduce significantly the state space, which leads to a numerically feasible analysis. Finally, quantum walks with particular boundaries can be used to simulate various (finite-size) physical systems.

First, we will consider boundary conditions for particles on a line. Then, we turn to the case of the square lattice. The correspondence between the two cases is given by (4) and (5).

For a particle on a line, two simplest cases are circular and the reflecting boundary conditions. In the case of circular boundary conditions, we connect two points M and $-M$. Our system is finite (a circle with $2M+1$ points), which affects the shift operator: the sum in (2) goes from $-(M-1)$ to $(M-1)$, while the connected points $-M$ and M are represented by adding the term

$$\begin{aligned}\hat{C} = & |-M\rangle\langle M| \otimes |R\rangle\langle R| + |M\rangle\langle -M| \otimes |L\rangle\langle L| \\ & + |M-1\rangle\langle M \otimes |L\rangle\langle L| + |-(M-1)\rangle\langle -M| \otimes |R\rangle\langle R|. \end{aligned} \tag{7}$$

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

For reflecting boundary conditions at positions M and $-M$, a coin operator that describes the reflection of particle's direction of movement (i.e. coin state) is changed at the points of reflection

$$(\hat{I}_P \otimes \hat{U}_C) \to \left(\sum_{x_1=-(M-1)}^{M-1} |x_1\rangle\langle x_1| \otimes \hat{U}_C \right) + (|-M\rangle\langle -M| + |M\rangle\langle M|) \otimes (|L\rangle\langle R| + |R\rangle\langle L). \quad (8)$$

Note that a general case of different two positions is equivalent to moving the initial position. One can also consider reflection in a single position M as well.

Combining the above two circular and reflecting (finite) together with open (infinite) boundary conditions, one can in the case of two particles on a line obtain different topologies: torus (two circles), finite and infinite cylinder (circular and reflection/open), square (both particles reflect on both sides of 0), etc. The corresponding corrected shift operators are obtained as a tensor product between the two single-particle corrected operators.

Analogously, one can study the circular and reflecting boundary conditions for the case of one particle on a two-dimensional lattice as well. For example, the correction of the shift operator for the case of circular conditions connecting points with different positions on the x-axis (M and $-M$), resulting in the (infinite) cylinder geometry, is given by

$$\hat{C}_x = \sum_y |-M, y\rangle\langle M, y| \otimes |E\rangle\langle E| + |M, y\rangle\langle -M, y| \otimes |W\rangle\langle W| \times |-(M-1), y\rangle\langle -M, y| \otimes |E\rangle\langle E| + |M-1, y\rangle\langle M, y| \otimes |W\rangle\langle W|, \quad (9)$$

while the reflection over the lines $x = \pm M$ results in the correction of the coin operator given by

$$\hat{R}_x = \sum_y |M, y\rangle\langle M, y| \otimes |W\rangle\langle E| + |-M, y\rangle\langle -M, y| \otimes |E\rangle\langle W| + \sum_{x=-(M-1)}^{M-1} \sum_{y=-M}^{M} |x, y\rangle\langle x, y| \otimes \hat{U}_C. \quad (10)$$

In addition to the above two cases, for a two-dimensional lattice one more option of a boundary along one axis occurs, presenting us with the Möbius strip. Connecting the points (M, y) of the line $x = M$, and $(-M, -y)$ of the line $x = -M$, results in the shift operator for which the sum for the x component is again going as above (from $-M+1$ to $M-1$), while the correction term is

$$\hat{M}_x = \sum_y |-M, -y\rangle\langle M, y| \otimes |E\rangle\langle E| + |M, -y\rangle\langle -M, y| \otimes |W\rangle\langle W| \times |M-1, y\rangle\langle M, y| \otimes |W\rangle\langle W| + |-(M-1), y\rangle\langle -M, y| \otimes |E\rangle\langle E|, \quad (11)$$

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

and analogously for connecting the $y = M$ and $y = -M$ axes according to Möbius topology. Combining the two results in the topology of Klein bottle.

Note that the Möbius boundary conditions for the case of two particles on a line would require a *nonlocal* shift operation, connecting *distant* sites x_2 and $-x_2$ of a second particle, whenever the first one is in the position $x_1 = M$ or $x_1 = -M$. Clearly, the shift operation is not a simple product of two one-particle operations. It is rather a controlled operation: conditioned to a position of the first particle, the second one is moving either locally ($x_2 \to j \pm 1$), if $x_1 \neq \pm M$, or nonlocally, otherwise. This can bring interesting consequences for the properties of a quantum walk, as it may introduce entanglement between the walkers.

Another type of boundary conditions are absorbing ones. They are modeled by measurements at certain points (of a line or a lattice) at each step of the walk. For example, an absorption at point x_1 on a line of a particle, coming from the left or right, is modeled by performing a measurement given by the projector $\hat{P}(x_1) = |x_1\rangle\langle x_1| \otimes \hat{I}_C$. If a particle is found at position i, it is absorbed and a walk stopped. Otherwise, a particle is for sure not in a position x_1, its state $|\psi(n)\rangle$, after n steps of the walk, is collapsed onto $\hat{I}_{PC} - \hat{P}(x_1) = (\hat{I}_P - |x_1\rangle\langle x_1|) \otimes \hat{I}_C|\psi(n)\rangle$, renormalized to unity, and evolved by the one-step evolution operator described above. Such walks, known as *measured walks*, were studied before in Refs. 2 and 24 in relation to various *hitting times* (see our discussion in the next section).

Partial measurements modeling absorption of a particle can be also used to model noise and decoherence effects. The other two ways to model noise and decoherence effects are breaking the links between certain nodes, or using different coins for different nodes and/or steps.

2.2. *Noise and decoherence: Broken links and different coins*

Studying decoherence effects by breaking the links between two nodes was first introduced in Ref. 25 for the case of a walk on a line, and later generalized to a two-dimensional case.[26] In a one-dimensional case, breaking the link between two neighboring nodes x_0 and $x_0 + 1$, at a step n of a walk, is equivalent to imposing the reflecting boundary conditions *in between* the two nodes, for particle coming from both directions. In other words, a particle coming from the node x_0 to the right will, instead of arriving to node $x_0 + 1$, change its direction (coin state, from $|R\rangle$ to $|L\rangle$) and return back to x_0, and analogously for the particle coming from the node $x_0 + 1$ to the left. Therefore, breaking the link between the nodes x_0 and $x_0 + 1$, at a step n of the walk, changes the shift operator $\hat{S}$ at that step from (2) to:

$$\hat{S} = \left(\sum_{x_1 \neq x_0} |x_1 + 1\rangle\langle x_1| \right) \otimes |R\rangle\langle R| + |x_0\rangle\langle x_0| \otimes |L\rangle\langle R| + \left(\sum_{x_1 \neq x_0 + 1} |x_1 - 1\rangle\langle x_1| \right) \otimes |L\rangle\langle L| + |x_0 + 1\rangle\langle x_0 + 1| \otimes |R\rangle\langle L|.$$

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

Analogously, the shift operator (6) for the case of a single walker in two dimensions is changed as well. For example, breaking the link between the nodes (x_0, y_0) and $(x_0 + 1, y_0)$ changes $|x_0 + 1, y_{)0}\langle x_0, y_0 \otimes |E\rangle\langle E|$ to $|x_0, y_0\rangle\langle x_0, y_0| \otimes |W\rangle\langle E|$, and $|x_0, y_0\rangle\langle x_0 + 1, y_0| \otimes |W\rangle\langle W|$ to $|x_0 + 1, y_0\rangle\langle x_0 + 1, y_0| \otimes |E\rangle\langle W|$.

Note that in the case of two walkers on a line, breaking of the links has to be identical for both walkers, if we want them to walk over the same line. Also, note that breaking of one link of only one of the two lines correspond to breaking of infinitely many links in the corresponding two-dimensional walk on the xOy plane.

If in every step different (possibly randomly chosen) links are broken, the shift operation changes from step to step and the breaking is dynamic. Breaking of the links can be either *static* or *dynamic*. If links (given by the pairs of nodes that are supposed to be connected by them) that are broken are fixed throughout the whole walk, the shift operator is changed according to the above description and *fixed* as well: in every step of the walk the same shift operation is applied (breaking is static). Otherwise, if in every step different (possibly randomly chosen) links are broken, the shift operation changes from step to step and the breaking is dynamic. Note that static breaking of the links allow to study walks over planar graphs with nodes having at most degree 4.

The other possibility of studying noise and decoherence effects is by introducing different coin operation for different nodes. In the one-dimensional case (analogously for two dimensions and higher), changing the unique coin operation $\hat{U}_C$ to $\hat{U}_{C_i}$, for each node i, changes the one-step evolution (1), from $\hat{U} = \hat{S}(\hat{I}_P \otimes \hat{U}_C)$ to

$$\begin{aligned}\hat{U} &= \hat{S}\left(\sum_{x_1} |x_1\rangle\langle x_1| \otimes \hat{U}_{C_{x_1}}\right) \\ &= \left(\sum_{x_1} |x_1 + 1\rangle\langle x_1 \otimes |R\rangle\langle R| \hat{U}_{C_{x_1}} + |x_1 - 1\rangle\langle x_1| \otimes |L\rangle\langle L| \hat{U}_{C_{x_1}}\right),\end{aligned}$$

and analogously for one particle on the lattice.

Again, altering coin operations can be either *static* or *dynamic*. Quantum-to classical transition driven by many coins was studied in Ref. 27. The particle localization due to varying coin in time was also studied in Refs. 21–23. The physical explanation of the particle localization can be explained as follows: a quantum walk, say on a lattice, can be seen as a model for a scattering process of a particle (say, electron) over the ions of a crystal lattice, such that a coin operators model the transition amplitudes between the neighboring ions. Introducing for some nodes coins different from the common one corresponds to introducing impurities (ions different from the one of a crystal lattice), which in the static case (impurities fixed in time) leads to Anderson-type particle localization.[28]

3. Quantities Computed by the Simulator

In this section, we describe the quantities that characterize quantum walk that our simulator is computing. Since the emphasis of the program is to study the joint

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

properties of a multi-particle quantum walk, we will use notation compatible with the case of two walkers on a line. Using 4 and 5, one can easily obtain the corresponding quantities for the case of a single walker in two dimensions.

A pure state of two walkers is a unit vector in a two-particle Hilbert space $\mathcal{H}_{12} \equiv \mathcal{H}_1 \otimes \mathcal{H}_2$, where one-particle mutually isomorphic spaces $\mathcal{H}_i = \mathcal{H}_{P,i} \otimes \mathcal{H}_{C,i}$, with $i \in \{1, 2\}$, have each a position and a coin factor space. Writing $\mathcal{H}_{12} = (\mathcal{H}_{P,1} \otimes \mathcal{H}_{P,2}) \otimes (\mathcal{H}_{C,2} \otimes \mathcal{H}_{C,2}) = \mathcal{H}_{P,12} \otimes \mathcal{H}_{C,12}$, one can talk of a two-particle position and coin (generally mixed) states.

If the initial state is $|\psi(0)_{12}\rangle \in \mathcal{H}_{12}$, then after n steps the state is

$$|\psi(n)_{12}\rangle = \hat{U}^n_{12}|\psi(0)\rangle_{12}.$$

Partial one-particle mixed states, after n steps of a walk, are given by density operators obtained by performing partial trace, $\hat{\rho}_1(n) = \mathrm{Tr}_2|\psi(n)\rangle\langle\psi(n)|_{12}$, and analogously for the second particle. The joint position state is evaluated by performing partial trace over the joint coin space $\mathcal{H}_{C,12}$, $\hat{\rho}_{P,12}(n) = \mathrm{Tr}_{C,12}|\psi(n)\rangle\langle\psi(n)|_{12}$, while to obtain the coin state, we do a partial trace over the joint position space $\mathcal{H}_{P,12}$, $\hat{\rho}_{C,12}(n) = \mathrm{Tr}_{P,12}|\psi(n)\rangle\langle\psi(n)|_{12}$. The one-particle position and coin states are obtained analogously, from one-particle states $\hat{\rho}_i(n)$, with $i \in \{1, 2\}$.

The main quantity from which we compute relevant joint properties is a joint two-particle probability distribution $p_{12}(i, j; n)$, a probability that the position x_1 of the first particle is i, and the position x_2 of the second is j. It is easily computed from the overall position state $\hat{\rho}_{P,12}(n)$ as

$$p_{12}(i, j; n) = {}_{12}\langle i, j|\hat{\rho}_{P,12}(n)|i, j\rangle_{12}. \tag{12}$$

Often, for reasons of simplicity, we assume the time dependence (i.e. step n) as implicit, and write $p_{12}(i, j)$, $\hat{\rho}_{P,12}$, etc. Also, sometimes we will drop the labels 12, 1 and 2 that denote whether a given quantity is a two-particle or a single-particle quantity: for example $p(i, j)$ instead of $p_{12}(i, j; n)$, or $\hat{\rho}_P$ instead of $\hat{\rho}_{P,12}(n)$.

Marginal probability, $p_1(i; n) = \sum_j p_{12}(i, j; n)$ can also be obtained from partial one-particle state, $p_1(i; n) = {}_1\langle i|\hat{\rho}_1(n)|i\rangle_1$, and analogously for particle 2.

In Ref. 12, it was shown that joint probability distributions obtained from initial states entangled in the coin degree of freedom cannot be achieved by initial separable states, while the marginal probabilities are nothing but a weighted sum of single-particle probability distributions. As this is the key result showing the difference between classical and quantum states at the level of joint, rather than single-walker level, we present it as a theorem (as well as the results regarding the average squared distance and the one-shot probability to hit, presented below).

Theorem 1. *Let $p^{\pm}_{12}(i, j; n)$ be the joint probability distributions obtained after n steps from the initial states $|\psi(0)\rangle^{\pm}_{12} = |00\rangle(|RL\rangle \pm |LR\rangle)/\sqrt{2}$ maximally entangled in the coin degree of freedom, starting from the initial position $(0, 0)$. Then, no initial state separable in the coin degree of freedom can produce either of the two joint probabilities $p^{\pm}_{12}(i, j; n)$.*

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

As the evolution of quantum walks in unitary, there exist no stationary state of the system, a fixed point of the evolution, as in the case of classical random walks. Therefore, there exist no stationary probability distribution (for detailed discussion, see for example Refs. 15 or 16). Yet, the time (or rather, step) average of the probability distribution, $\overline{p}_{12}(i,j;n) = \frac{1}{n}\sum_{k=1}^{n} p_{12}(i,j;k)$, does converge to a limiting distribution (analogously for one-particle distribution)

$$\pi_{12}(i,j) = \lim_{n\to\infty} \overline{p}_{12}(i,j;n).$$

One can thus study how fast (in steps n) an average distribution $\overline{p}_{12}(i,j;n)$ approaches the limiting one $\pi_{12}(i,j)$, globally (mixing time), point-wise (sampling time), etc. (for definitions of the mixing, sampling, filling and dispersion time, see for example Ref. 16).

The first global quantity of the two walker to consider is the average distance between the two

$$\langle \hat{d} \rangle = \langle |\hat{x}_1 - \hat{x}_2| \rangle = \sum_{i,j} p_{12}(i,j;n)|i-j|.$$

Note that the n-dependence is implicit as the average is an expectation value of an operator $\hat{d} = |\hat{x}_1 - \hat{x}_2|$ taken for the state $|\psi(n)\rangle_{12}$. This is clearly a global quantity that is not dependent only on marginal probability distributions $p_1(i;n)$ and $p_2(i;n)$, but on the correlations between the two random variables x_1 and x_2, which can in this case have particular entanglement-induced quantum features different from any classical-like correlations, analogously as in the case of the average squared distance $\langle \hat{d}^2 \rangle = \langle (\hat{x}_1 - \hat{x}_2)^2 \rangle$, as first shown in Ref. 12. We present this result as a theorem.

Theorem 2. *Let $\langle \hat{d}^2 \rangle^{\pm}$ be the average squared distance obtained from the initial states $|\psi(0)\rangle_{12}^{\pm} = |00\rangle(|RL\rangle \pm |LR\rangle)/\sqrt{2}$. Then, no initial state separable in the coin degree of freedom can produce either of the two expected values. Moreover, if by $\langle \hat{d}^2 \rangle^{\mathcal{S}}$ we denote the average squared distance of a generic initial state separable in the coin degree of freedom, then we have the following strict inequalities*:

$$\langle \hat{d}^2 \rangle^{+} < \langle \hat{d}^2 \rangle^{\mathcal{S}} < \langle \hat{d}^2 \rangle^{-}.$$

Next, we discuss various types of the so-called *mixing times*. First, we describe the one-particle case, then we move to the case of two walkers.

In algorithmic applications of quantum walks, like in search problems, a solution to the problem is given by a particular node i_0 (or more than one node, but the generalization is straightforward) and one is interested in the probability of finding this solution. In other words, we are interested in the probability of finding the walker in the position i_0.

Let us define two (one-particle) orthogonal projectors, $\hat{P}_0 = |i_0\rangle\langle i_0| \otimes \hat{I}_c$ and $\hat{P}_1 = \hat{I} - \hat{P}_0$, where $\hat{I}_c$ is the identity in the coin space. Then, the *one-shot hitting*

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

time for a given probability p is the number of steps $\mathcal{N}_o^{(1)}(i_0, p)$ for which the probability of the walker to be found in position i_0 is bigger or equal than p (see Ref. 3). It is determined by the (one-particle) *one-shot probability to hit*

$$\mathcal{P}_o^{(1)}(i_0; n) = \|\hat{P}_0|\psi(n)\rangle\|^2 = \|\langle i_0|\psi(n)\rangle\|^2. \tag{13}$$

The above definition of hitting time is useful in cases one can estimate step n around which the one-shot probability to hit $\mathcal{P}_o^1(i_0; n)$ is relatively high, like it was the case of Shor's algorithm.[29]

If we check after each step whether the particle is in position i_0, we effectively perform the above measurement $\hat{M}(i_0) = 0 \cdot \hat{P}_0 + 1 \cdot \hat{P}_1$ after each step. Such a walk, in which after each step the measurement given by $\hat{M}(i_0)$ is performed, is called the $|i_0\rangle$-*measured walk*[3]: if the particle is collapsed (absorbed) onto the ray $|i_0\rangle\langle i_0|$ (result 0 obtained), the solution to the problem (data-base search, etc.) is found and the walk terminated; otherwise, if the result is 1, the walk is evolved one more step by $\hat{U}$.

For $|i_0\rangle$-measured walk, let $\mathcal{P}_f^{(1)}(i_0; n)$ be the probability to detect (for the first time) the particle at position i_0, at step n (*first-time probability to hit*)

$$\mathcal{P}_f^{(1)}(i_0; n) = \|\hat{P}_0\hat{U}[\hat{P}_1\hat{U}]^{n-1}|\psi(0)\rangle\|^2. \tag{14}$$

Then, the *average hitting time* of the $|i_0\rangle$-measured walk is[24]

$$\mathcal{N}_a^{(1)}(i_0) = \sum_{n=1}^{\infty} n\mathcal{P}_f^{(1)}(i_0; n). \tag{15}$$

The above hitting time corresponds to a typical (average, expected) running time for the quantum-walk based algorithm.[24]

Finally, one might be interested in a number of steps $\mathcal{N}_c^{(1)}(i_0, p)$ after which a $|i_0\rangle$-measured walk has probability to stop greater or equal than certain given p. Such $\mathcal{N}_c^{(1)}(i_0, p)$ is called the *concurrent hitting time*,[3] and is given by the *concurrent probability to hit* $\mathcal{P}_c^{(1)}(i_0; n)$ for the walk to stop at any of the steps $n' \leq n$:

$$\mathcal{P}_c^{(1)}(i_0; n) = \sum_{n'=1}^{n} \|\hat{P}_0\hat{U}[\hat{P}_1\hat{U}]^{n'-1}|\psi(0)\rangle\|^2. \tag{16}$$

As noted in Ref. 24, the concurrent hitting time corresponds to the number of steps after which the probability to find the solution is greater or equal than p.

The above hitting times were defined for one-particle quantum walks and are given by the probabilities (13)–(16). For two particles, the corresponding hitting times are given by the analogous two-particle probabilities $\mathcal{P}_o^{(2)}(i_0; n)$, $\mathcal{P}_f^{(2)}(i_0; n)$ and $\mathcal{P}_c^{(2)}(i_0; n)$ that *at least one* of the two particles is detected in position i_0. Indeed, if the solution to the problem is given by the marked position i_0, it is enough if only one of the two walkers finds it. The two-particle probabilities and hitting times are obtained as in the case of one-particle walks, by substituting one-particle $\hat{P}_0$ (and its

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

complement $\hat{P}_1$) by its two-particle equivalent

$$\hat{P}_0 = (|i_0\rangle\langle i_0|_{P,1} \otimes \hat{I}_{P,2} + \hat{I}_{P,1} \otimes |i_0\rangle\langle i_0|_{P,2} - |i_0\rangle\langle i_0|_{P,1} \otimes |i_0\rangle\langle i_0|_{P,2}) \otimes \hat{I}_{C,12}$$

in Eqs. (13)–(16). Note the subtraction of the term $|i_0\rangle\langle i_0|_{P,1} \otimes |i_0\rangle\langle i_0|_{P,2}$ which is twice counted in the sum of the first two terms of the expression. Formally, without the subtraction this would not be an idempotent projector for which $\hat{P}_0^2 = \hat{P}_0$.

In finding solutions by searching certain data base, for example, we would prefer if the two particles search different regions at each given moment: if the two walkers are following each other, then there is little help of such "parallel processing". In other words, the probability that both walkers are found in the same position should be as small as possible; their average distance as big as possible. In Ref. 12, it was shown that the maximal/minimal one-shot probability $\mathcal{P}_o^2(i_0; n)$ to find at least one particle in position i_0 corresponds to the case of maximal/minimal average distance, when the initial coin state is the maximally entangled. This is a content of the following theorem.

Theorem 3. *Let $\mathcal{P}_o^2(i_0; n)^{\pm}$ be the average squared distance obtained from the initial states $|\psi(0)\rangle_{12}^{\pm} = |00\rangle(|RL\rangle \pm |LR\rangle)/\sqrt{2}$. Then we have the following strict inequalities*:

$$\mathcal{P}_o^2(i_0; n)^+ < \mathcal{P}_o^2(i_0; n)^{\mathcal{S}} < \mathcal{P}_o^2(i_0; n)^-,$$

where $\mathcal{P}_o^2(i_0; n)^{\mathcal{S}}$ is the one-shot probability of a generic initial state separable in the coin degree of freedom.

Global quantities do not depend on marginal probability distributions, but on the *correlations* between the random variables, in our case the positions of the two walkers. They are given by the covariance. For the case of the positions x_1 and x_2 of the two walkers, the covariance is given by

$$\mathrm{Cov}(x_1, x_2) = \langle (x_1 - \langle x_1\rangle)(x_2 - \langle x_2\rangle)\rangle = \langle x_1 x_2\rangle - \langle x_1\rangle\langle x_2\rangle, \tag{17}$$

The other way to quantify correlations between two walkers is by classical (Shannon) and quantum (von Neumann) mutual information. Classical (Shannon) mutual information between the two random variables, say positions x_1 and x_2, is (note the implicit dependence on the step n)

$$\mathcal{I}(x_1 : x_2) = H(x_1) + H(x_2) - H(x_1, x_2), \tag{18}$$

where $H(x_p) = -\sum_i p_1(i) \log p_1(i)$, with $p = 1, 2$, is the Shannon entropy of the random variable x_p taking the values $i \in \mathbb{Z}$, and

$$H(x_1, x_2) = -\sum_{i,j} p_{12}(i, j) \log p_{12}(i, j) \tag{19}$$

is the joint (Shannon) entropy of x_1 and x_2.

The corresponding quantum (von Neumann) mutual information between the position degrees of freedom of two walkers is given in terms of their global and partial

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

position states $\hat{\rho}_{P,12}$, $\hat{\rho}_{P,1}$ and $\hat{\rho}_{P,2}$

$$\mathcal{I}(\hat{\rho}_{P,12}) = S(\hat{\rho}_{P,1}) + S(\hat{\rho}_{P,2}) - S(\hat{\rho}_{P,12}), \tag{20}$$

where $S(\hat{\rho}_{P,1}) = -\mathrm{Tr}(\hat{\rho}_{P,1} \log \hat{\rho}_{P,1})$, and analogously for other two mixed states.

As it was shown in Ref. 12, the initial entanglement in the joint two-particle coin state can, starting from the initially product position state $|0,0\rangle_{12}$, bring about correlations between the positions of the two, the correlations beyond those achievable by any classical (i.e. mixed, but separable) initial state. In other words, a two-particle quantum walk can, in the course of steps, transfer the entanglement, and thus correlations, from the coin to the position degree of freedom. Therefore, one might be interested in analyzing the dynamics of mutual information (Shannon and von Neumann) between the (joint) coin and position degrees of freedom, or between the coins of the two walkers, or finally between the two walkers. They are given by expressions analogous to (18) and (20).

Finally, one can directly study entanglement between the two degrees of freedom. This is, being purely quantum feature, the most interesting quantity to study since it brings, in some settings, features of quantum walks classically impossible to achieve.[12] Nevertheless, unlike the correlations or (classical and quantum) mutual information, entanglement is more complex to characterize and quantify. In the case of pure bipartite states, entanglement between the two degrees of freedom is well defined and easy to evaluate: it is nothing but the von Neumann entropy of either of the two partial mixed states. Thus, if the initial state is pure, entanglement between the coin and position degree of freedom is given by

$$E_{C,P} = S(\hat{\rho}_{C,12}) = S(\hat{\rho}_{P,12}). \tag{21}$$

Obviously, from the numerical point of view, it is much easier, and more accurate, to use the first equality and deal with much smaller 4×4 matrix representation of $\hat{\rho}_{C,12}$. Writing the joint two-particle state $|\psi\rangle$ in the Schmidt bi-orthogonal expansion between the joint position and coin degrees of freedom

$$|\psi\rangle = \sum_{k=1}^{4} \sqrt{r_k} |\varphi_k\rangle_P |k\rangle_C,$$

the partial coin and position mixed states are given as $\hat{\rho}_C = \sum_{k=1}^{4} r_k |k\rangle\langle k|_C$ and $\hat{\rho}_P = \sum_{k=1}^{4} r_k |\varphi_k\rangle\langle\varphi_k|_P$ (note that, for reasons of simplicity, we dropped the step n dependence, and subscripts 12).

Calculating the coin state is computationally easy, its complexity is only quadratic in the number of steps n, as one has to evaluate $|\psi(n)\rangle$ and then a simple trace $\hat{\rho}_C(n) = \mathrm{Tr}_P |\psi(n)\rangle\langle\psi(n)|$. Solving the characteristic equation and finding eigenvalues r_k and eigenvectors $|k\rangle_c$ of a four-dimensional system given by $\hat{\rho}_C$ is easy task as well. Finally, the eigenvectors $|\varphi_k\rangle_P$ are easily calculated by obtaining the partial scalar product, $|\varphi_k\rangle_P = \frac{1}{\sqrt{r_k}} \langle k|\psi\rangle$. Finally, the partial position state is $\hat{\rho}_P = \sum_{k=1}^{4} r_k |\varphi_k\rangle\langle\varphi_k|_P$ and the entanglement between the position and the coin degree of

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

freedom is

$$E_{C,P} = S(\hat{\rho}_{C,12}) = S(\hat{\rho}_{P,12}) = -\sum_{k=1}^{4} r_k \log(r_k).$$

But, finding the entanglement between the positions of two walkers, or the two coins, is much more difficult problem. The partial position and coin states, $\hat{\rho}_{P,12}$ and $\hat{\rho}_{C,12}$, are mixed, and mixed-state entanglement is neither unique, nor easy to evaluate. One possible mixed-state entanglement measure is the *entanglement of formation.*[30] For the case of two-particle position state, one possible upper bound to the entanglement of formation could be given as

$$E_F(\hat{\rho}_{P,12}) = \sum_{k=1}^{4} r_k E(|\varphi_k\rangle_{P,12}),$$

with $E(|\varphi_k\rangle_{P,12}) = S(\mathrm{Tr}_{P,2}|\varphi_k\rangle\langle\varphi_k|_{P,12})$.

A relevant measure for the quantumness of correlations is the quantum discord.[31] In classical information theory, we have that the mutual information between random variables X and Y is given by $\mathcal{I}(X:Y) = H(X) + H(Y) - H(X,Y)$ or, equivalently, $\mathcal{J}(X:Y) = H(Y) - H(Y|X)$.

For the general joint state $\hat{\rho}_{XY}$, a subsystem's partial state after a measurement performed on the other sub-system is determined by the choice of the measurement and its outcome. Let $\hat{M}_x = \sum_i i\hat{\Pi}_i^X$ be a subsystem's X observable, where $\hat{\Pi}_i^X = |i\rangle\langle i|$ represent one-dimensional orthogonal projectors corresponding to the measurement outcome i. Upon measuring $\hat{M}_X \otimes \hat{I}_Y$ onto the joint state $\hat{\rho}_{XY}$, and obtaining the result i, the partial state of the subsystem Y is

$$\hat{\rho}_{Y|\hat{\Pi}_i^X} = \mathrm{Tr}_X \frac{(\hat{\Pi}_X \otimes \hat{I}_Y)\hat{\rho}_{XY}(\hat{\Pi}_X \otimes \hat{I}_Y)}{p_i},$$

where $p_i = \mathrm{Tr}_{XY}(\hat{\rho}_{XY}(\hat{\Pi}_i^X \otimes \hat{I}_Y))$. The expected entropy of the subsystem Y, conditioned by the measurement $\hat{M}_X$ performed on the subsystem X is

$$S(Y|\hat{M}_i^X) = \sum_i p_i S(\hat{\rho}_{Y|\hat{\pi}_i^X}).$$

The difference between the uncertainty of the state of the subsystem Y (measured by the von-Neumann entropy) before the subsystem measurement $\hat{M}_x$ was performed, and the expected uncertainty after it has been performed is

$$\mathcal{J}(X:Y)_{\hat{M}_x} = S(\hat{\rho_Y}) - S(Y|\hat{M}_X),$$

where $\hat{\rho}_Y = \hat{\rho}_{XY}$. The quantum discord, with respect to the measurement $\hat{M}_X$, is given by

$$\begin{aligned}\delta(X:Y)_{\{\hat{\Pi}_i^X\}} &= \mathcal{I}(X:Y) - \mathcal{J}(X:Y)_{\hat{M}_x} \\ &= S(\hat{\rho}_Y) - S(\hat{\rho}_{XY}) + S(Y|\hat{M}^X). \end{aligned} \tag{22}$$

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

In the case of two particles on the line, X denotes the position of particle 1 and Y the position of particle 2. For one particle on the lattice, X and Y denote the positions of the particle along the x and y axes, respectively.

4. The Simulator at Work

Here, we illustrate the simulator at work. The simulator is constituted by three main programs: one for simulating one particle on a square lattice; another for simulating two particles on a line; and, finally, one to simulate a single particle on a line. The simulator can output all the quantities described in Sec. 3, and are chosen in a configuration file. The configuration file is parsed by the simulator and contains the following information: dimension of the grid, number of steps to simulate, initial state for the particles, broken links in the grid, boundary conditions of the grid and measurement points, distribution of random quantum coins over the grid, quantities to simulate.

Next, we present several examples of usage of the simulator, one for each program.

4.1. *A particle on a square lattice*

For this case, we illustrate the simulator by computing several relevant quantities for a quantum walk over a square lattices with several broken links. The idea is to get a picture of the effect of broken links (or impurities) in the simulated quantities. In particular, we verify Anderson-like localization by noting that the average probability distribution is concentrated around the initial position.

For this particular simulation, we consider:

- a grid of size 61×61;
- a number of steps 30;
- broken link probability of 0.48;
- initial state $|\psi(0)\rangle = |0,0\rangle|E\rangle$.

The plots output by the program concerning the quantities are depicted in Figs. 2–5.

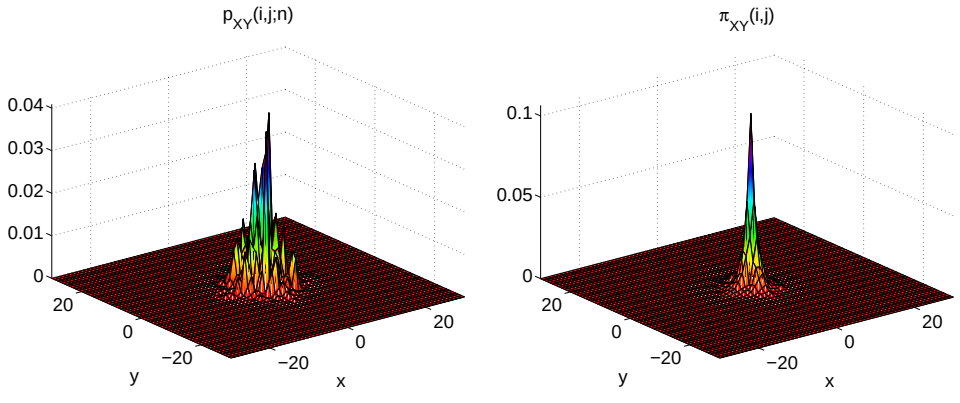
4.2. *Two particles on a line*

The effect of entanglement on quantum walks is a relevant problem which is hard to tackle analytically. For this reason, the simulator is designed to determine various information theoretical quantities, such as entropy, mutual information and (upper-bound) entanglement of formation, which quantify the effect of entanglement between the coins on the joint position probability distribution of two walkers. Indeed, one can check that entangled particles evolve differently than non entangled ones.[12]

For this particular simulation, we consider

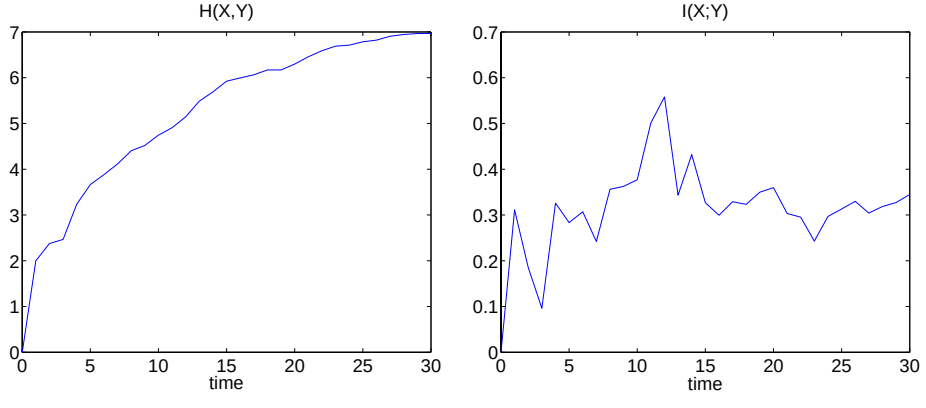
- a grid of size 61×61;
- a number of steps 30;

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.



(a) (b)

Fig. 2. Evolution of one particle on a square lattice with broken link probability of 0.48 for 30 steps. (a) Probability on plane after 30 steps. (b) Average probability distribution.



(a) (b)

Fig. 3. Evolution of one particle on a square lattice with broken link probability of 0.48 for 30 steps. (a) Shannon entropy for variables X and Y. (b) Shannon mutual information for position variables X and Y.

- initial state $|\psi(0)\rangle_{12} = |0,0\rangle_{12}(|RR\rangle - |LL\rangle)$.

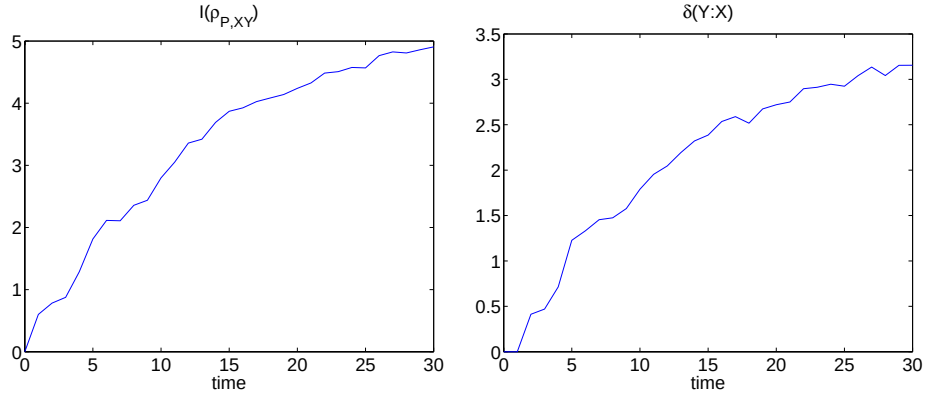
The results of the quantities are depicted in Figs. 6–9.

4.3. *A particle on a line*

Although the case of a particle on a line is quite well studied, we also included it in the simulator. To take profit of the features of the simulator, we considered an absorbing boundary condition. For this particular simulation, we consider

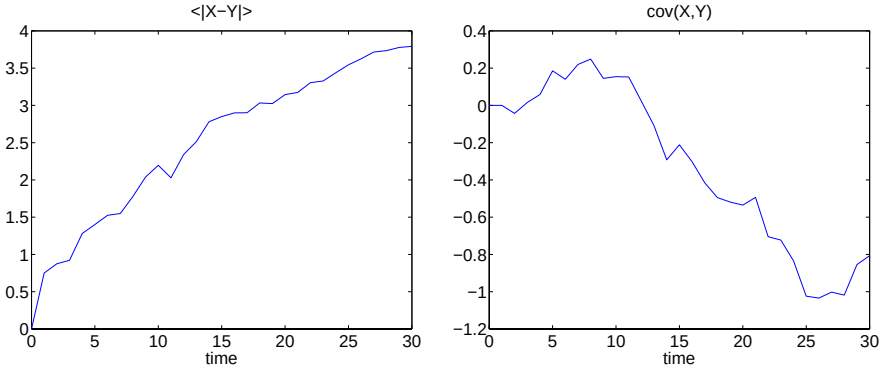
- a line of length 100;

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.



(a) (b)

Fig. 4. Evolution of one particle on a square lattice with broken link probability of 0.48 for 30 steps. (a) von Neumann mutual information of $\hat{\rho}_{P,XY}$. (b) Quantum discord of Y given measurements $\hat{\Pi}_i^X$ on X.



(a) (b)

Fig. 5. Evolution of one particle on a square lattice with broken link probability of 0.48 for 30 steps. (a) Mean distance of variables X and Y. (b) Covariance of variables X and Y.

- a number of steps 10000;
- absorbing boundary condition;
- initial state $|\phi_0\rangle = |0\rangle|R\rangle$.

The results of the quantities are depicted in Fig. 10.

Compared to the Discrete Time Quantum Walks (DTQW) in Ref. 32, our program can perform simulation of a walk for up to 10^7 steps, which might seem as a disadvantage. Nevertheless, the two types of quantum walks differ significantly in their long-time behavior, as the cited paper thoroughly analyses: while the standard DTQW has a ballistic behavior, i.e. its diffusion scales as $\sigma \sim t$, the feed-forward DTQW diffusion scales as $\sigma \sim t^{0.4}$. Consequently, the memory needed to encode the

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

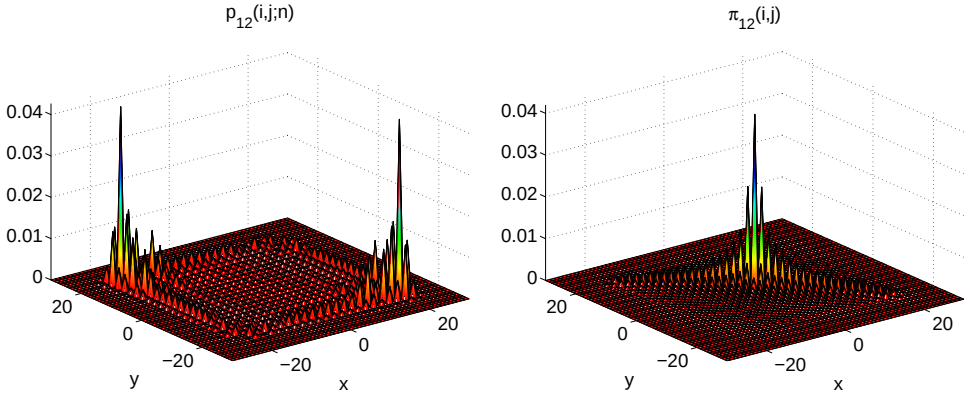


Fig. 6. Evolution of one particle on a square lattice with broken link probability of 0.48 for 30 steps. (a) Probability on plane after 30 steps. (b) Average probability distribution.

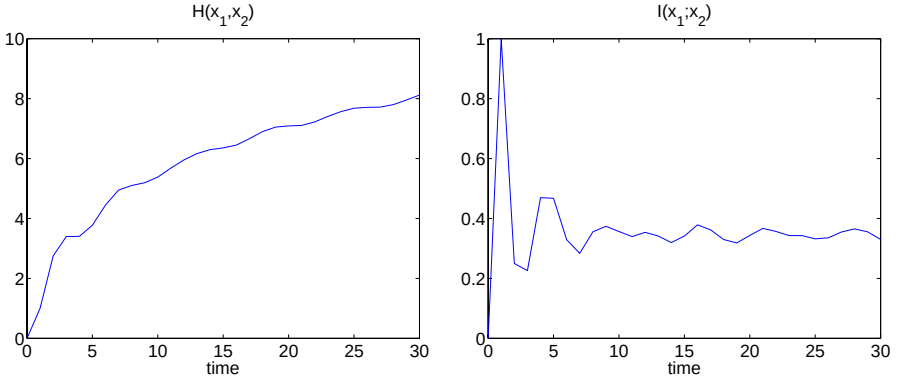
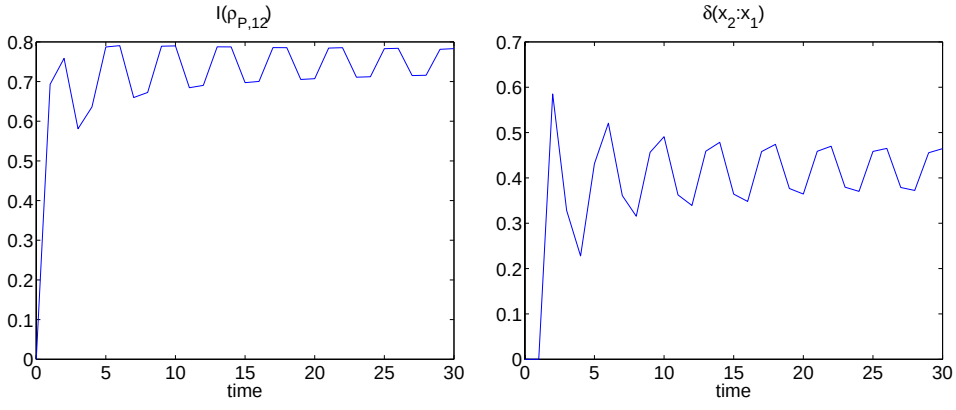


Fig. 7. Evolution of one particle on a square lattice with broken link probability of 0.48 for 30 steps. (a) Shannon entropy for variables x_1 and x_2. (b) Shannon mutual information for position variables x_1 and x_2.

relevant part of the state is just a fraction of the memory needed to encode a state of a standard one-dimensional DTQW after 10^8 steps (see Fig. 2.A of Ref. 32, where the x-axis goes from -4500 to $+4500$). Indeed, the memory needed to encode a quantum state of a standard DTQW after 10^8 steps would be 12 800 Gbits, which exceeds the limits of any conceivable computer memory that can be found on the market.

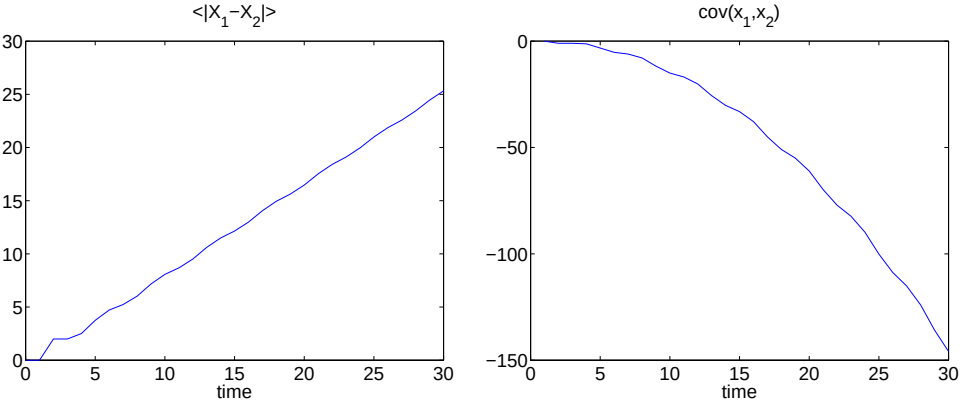
Nevertheless, when posing similar constraints to a standard one-dimensional DTQW, for example reflecting boundary conditions, thus effectively limiting the area of a walk, our simulator can perform many more steps. In particular, for a line that goes from node -4000 to node $+4000$, our program was able to finish the simulation for 10^6 steps in roughly 3 h, while to simulate 10^7 steps took roughly 30 h.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.



(a) (b)

Fig. 8. Evolution of one particle on a square lattice with broken link probability of 0.48 for 30 steps. (a) von Neumann mutual information of $\hat{\rho}_{P,12}$. (b) Quantum discord.



(a) (b)

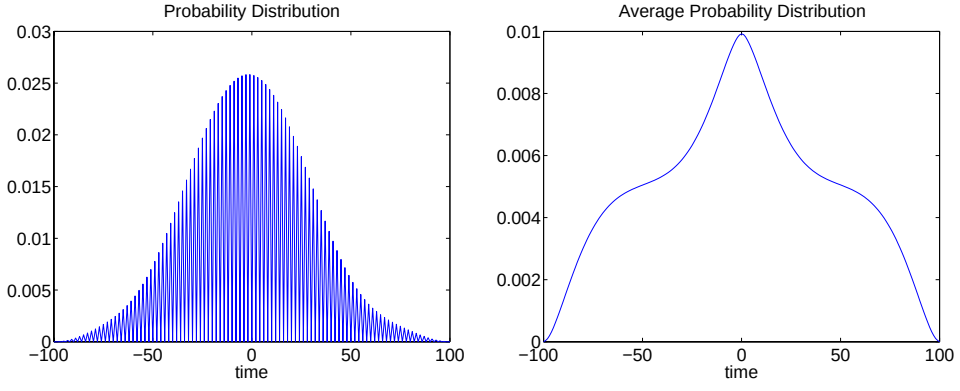
Fig. 9. Evolution of one particle on a square lattice with broken link probability of 0.48 for 30 steps. (a) Mean distance of variables x_1 and x_2. (b) Covariance of variables x_1 and x_2.

Consequently, running time for 10^8 steps would take roughly about two weeks, which is reasonable for a well-defined scientific research. In Fig. 11, we present the probability distribution and the average probability distribution of particle position for the mentioned simulation of 10^7 steps.

4.4. *Example*: *Anderson localization*

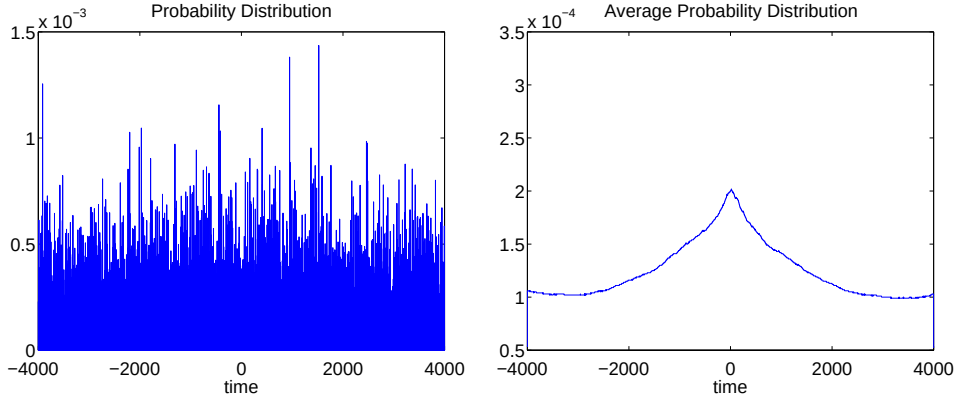
Here we simulate the effect of two types of decoherence due to random broken links and random coins, in order to see if Anderson localization will occur, as well as to find

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.



(a) (b)

Fig. 10. Evolution of one particle on a line of length 201, for 10 000 steps with absorbing boundary. (a) Probability distribution for position. (b) Average probability distribution for position.



(a) (b)

Fig. 11. Probability distribution (a) and average probability distribution (b) of the particle position for one-particle quantum walk on the open line after 10^7 steps. The initial state is $|\psi(0)\rangle = |0\rangle|L\rangle$, and the random coin parameters are set within the interval $\theta, \zeta, \xi \in [\frac{\pi}{4} - \frac{\pi}{8}, \frac{\pi}{4} + \frac{\pi}{8}]$.

some difference between the two models. In the case of random broken link, the broken link factor gives the probability of each link to be broken. For random coins, the random coin factor is the probability of the coin operator, at each position, to be randomly chosen.

For this particular simulation, we consider

- a two-dimensional lattice of size 61×61;
- a number of steps 100;
- Klein bottle boundary condition;
- initial state $|\phi_0\rangle = |0\rangle(|E\rangle + |S\rangle)$.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

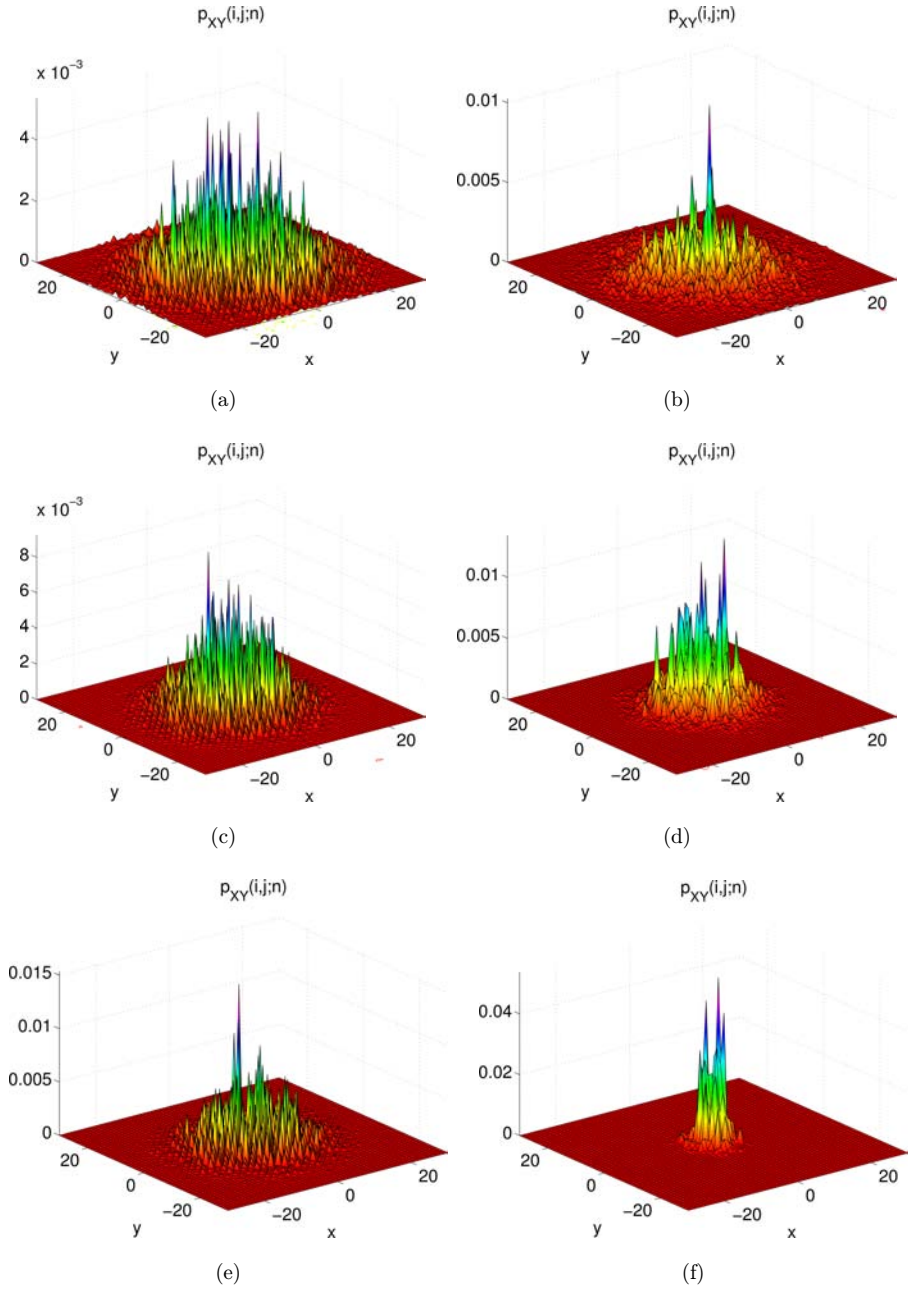


Fig. 12. Position probability distribution of one particle on the lattice of dimension 61×61, for 100 steps with different random coin factors and random broken link factors. (a) Probability distribution for position with random coin factor 0.2. (b) Probability distribution for position with random broken link factor 0.2. (c) Probability distribution for position with random coin factor 0.48. (d) Probability distribution for position with random broken link factor 0.48. (e) Probability distribution for position with random coin factor 0.9. (f) Probability distribution for position with random broken link factor 0.9.

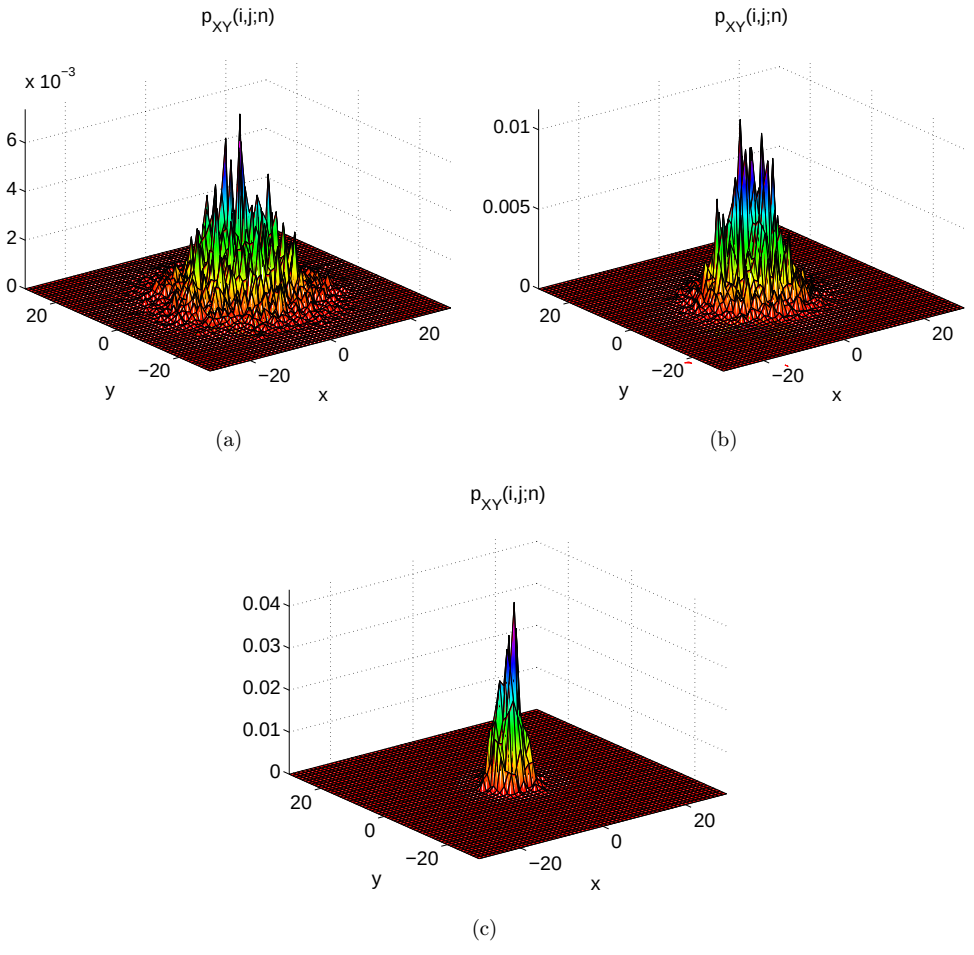


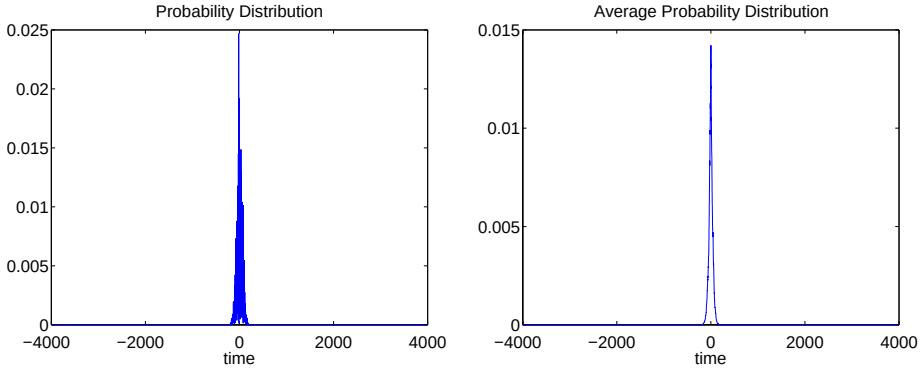
Fig. 13. Position probability distribution of particle on the lattice of dimension 61×61, for 100 steps with same both broken link and random coin factors. (a) Position probability distribution with broken link and random coin factors 0.2. (b) Position probability distribution with broken link and random coin factors 0.48. (c) Position probability distribution with broken link and random coin factors 0.9.

- Simulations ran for 9 cases: random coin factors of 0.2, 0.48 and 0.9, random broken link factors of 0.2, 0.48 and 0.9 and random coin and broken link factors of 0.2, 0.48 and 0.9.

The results of the quantities are depicted in Figs. 12 and 13.

Figures 12 and 13 suggest that both random broken link and random coin can lead to Anderson localization. Moreover, based on Fig. 12 one can conclude that for the same factors, the localization is sharper for random broken links compared to random coins. Figure 13 shows the Anderson localization due to both random broken links and coins. Further analysis/simulation is needed to to characterize Anderson localization.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.



(a) (b)

Fig. 14. (a) Probability distribution and (b) average probability distribution of the particle position for one-particle quantum walk on the open line after 4000 steps. The initial state is $|\psi(0)\rangle = |0\rangle|R\rangle$, the probability that at each step a link will be broken (index broken link) is 0.3 and the random coin parameters are set within the interval $\theta, \zeta, \xi \in [\frac{\pi}{4} - \frac{\pi}{8}, \frac{\pi}{4} + \frac{\pi}{8}]$.

Anderson localization for quantum walk on the line can be achieved by defining static random coins of parameters $\theta, \zeta, \xi \in [\frac{\pi}{4} - \frac{\pi}{8}, \frac{\pi}{4} + \frac{\pi}{8}]$, illustrated in Fig. 14 with the following conditions:

- a line of size 8001;
- a number of steps 4000;
- reflecting boundary condition;
- initial state $|0\rangle|R\rangle$;
- static random coin with parameters $\theta, \zeta, \xi \in [\frac{\pi}{4} - \frac{\pi}{8}, \frac{\pi}{4} + \frac{\pi}{8}]$;
- the probability that at each step a link will be broken (index broken link) is 0.3.

In the case of two particles on separate lines, we can simulate two different types of decoherences leading to localization. In Fig. 15, one particle is under the influence of static random coins, the other is other the influence of random coins at random positions.

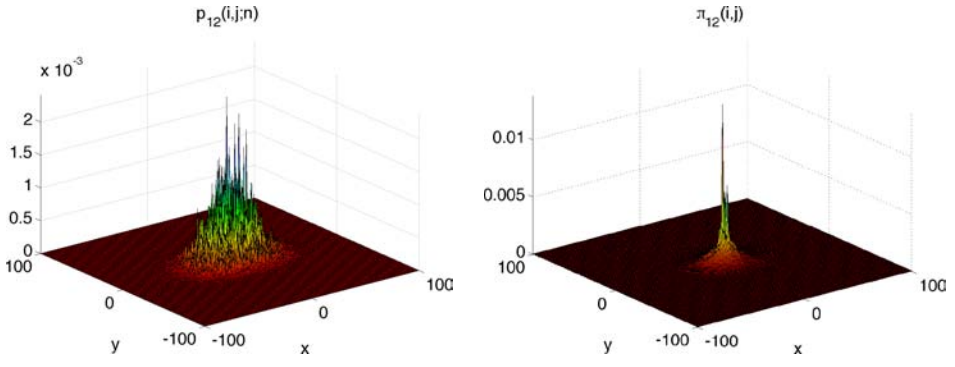
- two lines of size 201;
- a number of steps 100;
- reflecting boundary conditions;
- initial state $\sqrt{2}|0,0\rangle|RR\rangle$;
- static random coin with parameters $\theta, \zeta, \xi \in [\frac{\pi}{4} - \frac{\pi}{8}, \frac{\pi}{4} + \frac{\pi}{8}]$ for the first walker;
- broken link index of 0.3 for the second walker, with Hadamard coin.

4.5. *Example*: *Static broken links*

To illustrate the usage of static broken links, in Fig. 16 we consider a two-dimensional lattice with nine boxes of equal dimensions and slits between them, as indicated below:

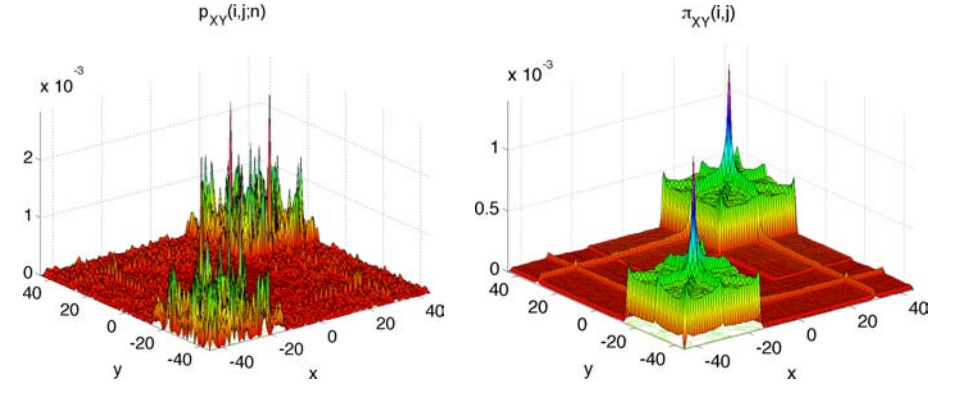
- two-dimensional lattice of size 91×91;

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.



(a) (b)

Fig. 15. (a) Probability distribution and (b) average probability distribution of the particles positions for two-particle quantum walk on open lines after 100 steps. The initial state is $|\psi(0)\rangle = |0,0\rangle|RR\rangle$, for the first walker the random coin parameters are set within the interval $\theta, \zeta, \xi \in [\frac{\pi}{4} - \frac{\pi}{8}, \frac{\pi}{4} + \frac{\pi}{8}]$, while for the second walker the fixed coin is given by the Hadamard operator and the index broken link is 0.3.



(a) (b)

Fig. 16. (a) Probability distribution and (b) average probability distribution of particle position for one-particle quantum walk on a lattice with reflecting boundary conditions at $x, y = \pm 45$ after 1 000 steps. The initial state is $|\psi(0)\rangle = \frac{1}{2}(|-30,-30\rangle(|E\rangle + i|N\rangle) + |30,30\rangle(|W\rangle + i|S\rangle))$, with the fixed coin given by the Hadamard operator. The static broken links are set between positions $(-15, y)$ and $(-14, y)$ and $(14, y)$ and $(15, y)$, for $y \in \{-45, \ldots, 45\} \setminus \{-30, 0, 30\}$, and positions $(x, -15)$ and $(x, -14)$ and $(x, 14)$ and $(x, 15)$, for $x \in \{-45, \ldots, 45\} \setminus \{-30, 0, 30\}$. Note that the entire grid is divided into nine equally-sized loosely connected squares, and the initial state of the walker is a linear superposition of two distant positions (and the corresponding coin states) located in different squares.

- a number of steps 1000;
- reflecting boundary conditions;
- initial state $\frac{1}{2}(|-30,-30\rangle(|E\rangle + i|N\rangle) + |30,30\rangle(|W\rangle + i|S\rangle))$;
- Hadamard coin;

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

- the static broken links are set between positions $(-15, y)$ and $(-14, y)$ and $(14, y)$ and $(15, y)$, for $y \in \{-45, \ldots, 45\} \setminus \{-30, 0, 30\}$, and positions $(x, -15)$ and $(x, -14)$ and $(x, 14)$ and $(x, 15)$, for $x \in \{-45, \ldots, 45\} \setminus \{-30, 0, 30\}$;

5. Conclusion

We developed a simulator for a two-particle quantum walk on a line and one particle on a two-dimensional squared lattice. The simulator can be used to investigate the equivalence between the two cases (one- and two-particle walks) for various boundary conditions (open, circular, reflecting, absorbing and their combinations). For the case of a single walker on a two-dimensional lattice, the simulator can implement the Möbius strip and other similar boundary conditions. Furthermore, other topologies for the walker are also simulated by the proposed tool, like planar graphs with degree up to 4, by considering missing links over the lattice. The simulator is able to compute a vast number of relevant quantities, namely: average position probability distribution, standard deviation/covariances, Shannon and von Neumann entropy and mutual information, upper bound for the entanglement formation, as well as quantum discord. The simulator is available at http://qwsim.weebly.com/ and allows for computational experiments for quantum walks covering new aspects and quantities that were not available before.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

Acknowledgments

The authors acknowledge the security and Quantum Information Group of IT and IT project QbigD funded by FCT PEst-OE/EEI/LA0008/2013 and UID/EEA/50008/2013. JR thanks FCT PhD Grant SFRH/BD/75085/2010. PM also acknowledge the confident project PTDC/EEI-CTP/4503/2014.

References

1. Y. Aharonov, L. Davidovicg and N. Zagury, *Phys. Rev. A* **48**, 1687 (1993).
2. A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann and D. A. Spielman, Exponential algorithmic speedup by quantum walk, *Proc. 35th ACM Symp. Theory of Computing (STOC 2003)*, pp. 59–68, arXiv:quant-ph/0209131v2.
3. J. Kempe, *Probab. Theory Relat. Fields* **133**(2), 215 (2005), arXiv:quant-ph/0205083v1.
4. E. Farhi and S. Gutmann, *Phys. Rev. A* **58**, 915928 (1998).
5. A. M. Childs and J. Goldstone, *Phys. Rev. A* **70**, 022314 (2004).
6. S. Aaronson and A. Ambainis, arXiv:quant-ph/0303041v3.
7. N. Shenvi, J. Kempe and K. B. Whaley, *Phys. Rev. A* **67**, 052307 (2003).
8. A. Ambainis, J. Kempe and A. Rivosh, Coins make quantum walks faster, *SODA '05: Proc. Sixteenth Annual ACM-SIAM Symp. Discrete Algorithms*, p. 1099 (2005).
9. A. Tulsi, *Phys. Rev. A* **78**, 012310 (2008).
10. M. Szegedy, Quantum Speed-up of Markov Chain Based Algorithms, *Proc. 45th Annual IEEE Symp. Foundations of Computer Science*, p. 3241 (2004).
11. F. Magniez, A. Nayak, J. Roland, M. Santha, *SIAM J. Comput.* **40**(1), 142 (2011), arXiv:quant-ph/0608026v4.

12. Y. Omar, N. Paunković, L. Sheridan and S. Bose, *Phys. Rev. A* **74**, 042304 (2006).
13. M. Štefaňák, T. Kiss, I. Jex and B. Mohring, *J. Phys. A, Math. Gen.* **39**, 14965 (2006).
14. S. E. Venegas-Andraca and S. Bose, arXiv:0901.3946.
15. J. Kempe, *Cont. Phys.* **44**, 307 (2003), arXiv:quant-ph/0303081.
16. D. Aharonov, A. Ambainis, J. Kempe and U. Vazirani, Quantum Walks on graphs, *Proc. ACM Symp. Theory of Computation (STOC'01)* (2001), pp. 50–59, arXiv:quant-ph/0012090v2.
17. F. L. Marquezino and R. Portugal, *Comput. Phys. Commun.* **179**, 359 (2008).
18. S. Berry, P. Bourke and J. Wang, *Comput. Phys. Communi.* **182**(10), 2295 (2011).
19. I. Carneiro, M. Loo, X. Xu, M. Girerd, V. Kendon and P. L. Knight, *New J. Phys.* **7**, 156 (2005).
20. M. Annabestani, M. R. Abolhasani and G. Abal, *J. Phys. A, Math. Theor.* **43**, 075301 (2010).
21. M. C. Bañuls, C. Navarrete, A. Pérez, E. Roldán and J. C. Soriano, *Phys. Rev. A* **73**, 062304 (2006).
22. C. M. Chandrashekar, *Phys. Rev. A* **83**, 022320 (2011).
23. H. Obuse and N. Kawakami, *Phys. Rev. B* **84**, 195139 (2011).
24. H. Krovi and T. Brun, *Phys. Rev. A* **73**, 032341 (2006).
25. A. Romanelli, R. Siri, G. Abal, A. Auyuanet and R. Donangelo, *Phys. A* **347C**, 137 (2005).
26. A. C. Oliveira, R. Portugal and R. Donangelo, *Phys. Rev. A* **74**, 012312 (2006).
27. T. A. Brun, H. A. Carteret and A. Ambainis, *Phys. Rev. A* **67**, 052317 (2003).
28. P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
29. P. W. Shor, Algorithms for quantum computation: Discrete logarithms and factoring, *Proc. 35th Annual Symp. Foundations of Computer Science*, p. 124 (1994).
30. C. H. Bennett, D. P. DiVincenzo, J. Smolin and W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1996).
31. H. Ollivier and W. H. Zurek, *Phys. Rev. Lett.* **88**, 017901 (2001).
32. Y. Shikano, T. Wada and J. Horikawa, *Sci. Rep.* **4**, 442 (2014), doi: 10.1038/srep04427.

Int. J. Mod. Phys. C Downloaded from www.worldscientific.com
by UNIVERSITY OF TORONTO on 04/18/17. For personal use only.

International Journal of Quantum Information
Vol. 13, No. 6 (2015) 1550051 (25 pages)
© World Scientific Publishing Company
DOI: 10.1142/S0219749915500513

World Scientific
www.worldscientific.com

Secure *N*-dimensional simultaneous dense coding and applications

H. Situ

Department of Computer Science,
Sun Yat-sen University, Guangzhou 510006, P. R. China

College of Mathematics and Informatics,
South China Agricultural University, Guangzhou 510642, P. R. China
situhaozhen@gmail.com

D. Qiu

Department of Computer Science,
Sun Yat-sen University, Guangzhou 510006, P. R. China

The Guangdong Key Laboratory of Information Security Technology,
Sun Yat-sen University, Guangzhou 510006, P. R. China
issqdw@mail.sysu.edu.cn

P. Mateus* and N. Paunković†

Departamento de Matemática,
Instituto Superior Técnico, Universidade de Lisboa,
SQIG — Security and Quantum Information Group,
Instituto de Telecomunicações,
Avenida Rovisco Pais 1049-001, Lisboa, Portugal
**pmat@math.ist.utl.pt*
†*npaunkovic@gmail.com*

Received 24 June 2015
Accepted 18 October 2015
Published 10 December 2015

Simultaneous dense coding (SDC) guarantees that Bob and Charlie simultaneously receive their respective information from Alice in their respective processes of dense coding. The idea is to use the so-called locking operation to "lock" the entanglement channels, thus requiring a joint unlocking operation by Bob and Charlie in order to simultaneously obtain the information sent by Alice. We present some new results on SDC: (1) We propose three SDC protocols, which use different N-dimensional entanglement (Bell state, W state and GHZ state). (2) Besides the quantum Fourier transform, two new locking operators are introduced (the double controlled-NOT operator and the SWAP operator). (3) In the case that spatially distant Bob and Charlie have to finalize the protocol by implementing the unlocking operation through communication,

† Corresponding author.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

we improve our protocol's fairness, with respect to Bob and Charlie, by implementing the unlocking operation in series of steps. (4) We improve the security of SDC against the intercept–resend attack. (5) We show that SDC can be used to implement a fair contract signing protocol. (6) We also show that the N-dimensional quantum Fourier transform can act as the locking operator in simultaneous teleportation of N-level quantum systems.

Keywords: Quantum communication; teleportation; dense coding.

1. Introduction

Due to the Holevo bound,[1] at most $\log_2 N$ bits of information can be transmitted via a qu$N$it (N-level quantum system). Dense coding, proposed by Bennett and Wiesner[2] in 1992, increases the classical capacity of a quantum communication channel with the help of prior entanglement.[3] If the sender and the receiver share a pair of entangled qu$N$its, $2 \log_2 N$ bits can be transmitted via a qu$N$it.[2]

In the simplest case of dense coding two parties, Alice and Bob, share a pair of entangled qubits (two-level quantum systems) in a Bell state. Alice first performs one of the four local unitary operations I, σ_x, $i\sigma_y$ and σ_z (where σ_j are the Pauli matrices) on her qubit to encode 2 bits of information, transforming the entangled pair into one of the four mutually orthogonal Bell states. Then Alice sends her qubit to Bob through a quantum channel. Bob is now able to measure both qubits in the Bell basis to obtain one of the four possible outcomes correlated with the operations performed by Alice. Thus, Alice can transmit 2 bits of information to Bob by manipulating and sending only one qubit.

Thus far, dense coding has been extensively studied in various ways. For example, dense coding that utilizes high-dimensional entangled states has been studied in Refs. 4–6, non-maximally entanglement channels in Refs. 7–19, while multipartite entanglement channels have also been considered in Refs. 20–27. Another generalization is to perform the communication task under the control of a third party, so-called controlled dense coding.[28–31]

Inspired by the simultaneous teleportation scheme proposed by Wang *et al.*,[32] we have proposed a simultaneous dense coding (SDC) scheme,[33] which guarantees that Bob and Charlie (the receivers) simultaneously receive their respective information from Alice (the sender) in their respective processes of dense coding. In this scheme, Alice first performs a locking operation to entangle the particles from two independent quantum entanglement channels, and therefore the receivers cannot obtain their respective information separately, before performing the unlocking operation together. The quantum Fourier transform and its inverse are used as the locking and unlocking operators, respectively.

Simultaneous dense coding may be relevant and useful in various applications in which Bob and Charlie can be either close or separated. If Alice has two different secrets, one for Bob and another for Charlie, she can utilize SDC to guarantee that Bob and Charlie simultaneously reveal their respective secrets. Bob does not know Charlie's secret and vice versa. For example, boss wants two employees to

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

simultaneously carry out two confidential commercial activities under the condition that the sensitive information of each activity is only revealed to whom is in charge of that activity.

There are other applications improving some models or tasks of quantum communication. In Sec. 5, we show that SDC can be used to implement a fair contract signing protocol[34] *between spatially distant Bob and Charlie. In this case, Bob and Charlie are separated so that there is a problem of how to implement the unlocking operation fairly. We discuss this problem in Sec. 3.*

In this paper, we give some new results on SDC and teleportation. The main improvements over earlier proposals are: (i) introduction of new locking operations (the double-controlled-NOT and the SWAP operators) for SDC schemes that were not used before, showing that the protocol can be achieved by using three different N-dimensional quantum states (Bell, W and GHZ states) that were not mentioned in that context in the literature before; (ii) designing a strategy which achieves the fairness of SDC, allowing for the protocol to be performed even when mutually mistrustful Bob and Charlie are situated on spatially distant locations, a feature which previous SDC proposals do not allow; (iii) designing a strategy which also achieves the security of SDC against intercept–resend attack, improving the security of SDC over the previous proposals; (iv) introducing simultaneous teleportation scheme for the transfer of arbitrary number of N-level quantum states. In addition to that, we proposed contract signing protocol based on a fair and secure SDC protocol. Having more options for achieving simultaneous transmission of (densely coded) information presents opportunities for designing SDC protocols with improved security features, and in addition provides wider range of possibly suitable experimental realizations. Indeed, achieving the fairness of SDC scheme, a feature not satisfied by previous SDC proposals, is only possible with the double-controlled-NOT and the SWAP locking operator. Also, the security of the protocol against intercept–resend attack is a novel feature of SDC schemes introduced in our paper.

The paper is organized as follows. In Sec. 2, we introduce the locking operators (i.e. the quantum Fourier transform, the double controlled-NOT operator and the SWAP operator) and propose three SDC protocols utilizing different N-dimensional quantum states (i.e. Bell states, W states, and GHZ states). In Sec. 3, we improve the fairness of the protocol by implementing the unlocking operation between spatially distant Bob and Charlie. In Sec. 4, we improve the security of SDC against the intercept–resend attack. In Sec. 5, we show that SDC can be used to implement a fair contract signing protocol. In Sec. 6, we show that the N-dimensional quantum Fourier transform can act as the locking operator in simultaneous teleportation of qu$N$its. A brief conclusion follows in Sec. 7.

2. *N*-Dimensional Simultaneous Dense Coding

A qu$N$it is an N-dimensional quantum system. States of N-dimensional quantum systems can be mapped onto $n = \log_2(N)$ qubits, and throughout the paper we will

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

assume that a qu$N$it is realized as an ordered array of qubits. We also assume that the dimension N of qu$N$its satisfies the requirement $N = 2^n$, with $n \in \mathbb{N}$. In other words, one qu$N$it consists of (or can be mapped to) n qubits.

In the task of N-dimensional SDC, Alice intends to use dense coding to send two cNits (arrays of classical bits) $b_1, b_2 \in \{0, 1, \ldots, N-1\}$ to Bob and two cNits $c_1, c_2 \in \{0, 1, \ldots, N-1\}$ to Charlie, under the condition that Bob and Charlie must collaborate to simultaneously find out what she sends.

In the following subsections, we propose three protocols, one using N-dimensional Bell states, the other using W states, and the third using GHZ states, as the entanglement channels, respectively. The idea behind these protocols is to perform the locking operator on Alice's qu$N$its before sending them to Bob and Charlie. After receiving Alice's qu$N$its, the states of Bob's subsystem and Charlie's subsystem are independent of (b_1, b_2) and (c_1, c_2), respectively, so that they know nothing about the encoded information. Only after performing the unlocking operator (the inverse of the locking operator) together, Bob and Charlie can obtain (b_1, b_2) and (c_1, c_2), respectively.

In this section, we assume that Bob and Charlie are at the same site. In order to implement the unlocking operator, they can input their particles into a physical device that can unlock the particles, and then get their respective output particles. The problem of how distant Bob and Charlie can perform the unlocking operator fairly is discussed in Sec. 3.

2.1. *DCNOT locking operator*

Before introducing the protocols, let us first have a look at the double controlled-NOT (DCNOT) operator, which is used as the locking operator. The DCNOT operator is composed of two controlled-NOT (CNOT) operators. The CNOT operator has two input qubits. If the first qubit (the control qubit) is in state $|1\rangle$, CNOT flips the state of the second qubit (the target qubit). If the control qubit is in state $|0\rangle$, CNOT does nothing to the target qubit ($|0\rangle$ and $|1\rangle$ are the basis vectors). Its action can be described as ($x, y = 0, 1$ are bit values):

$$|x\rangle|y\rangle \to |x\rangle|x \oplus y\rangle, \tag{1}$$

where $\oplus$ denotes bitwise addition modulo 2.

The DCNOT operator is formed by performing the first CNOT with the first qubit as the control and the second being the target, and then a second CNOT with inverse roles of the qubits (the first qubit as the target and the second being the control). Its action can be described as:

$$|x\rangle|y\rangle \to |y\rangle|x \oplus y\rangle, \tag{2}$$

and the inverse DCNOT operator can be described as:

$$|x\rangle|y\rangle \to |x \oplus y\rangle|x\rangle. \tag{3}$$

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

Our qu$N$its $|x\rangle_{A_1}|y\rangle_{A_2}$, with $x, y \in \{0, 1, \ldots, N-1\}$, are arrays of n qubits, and can be written as:

$$|x_1, \ldots, x_n\rangle_{A_{1,1}\ldots A_{1,n}}|y_1, \ldots, y_n\rangle_{A_{2,1}\ldots A_{2,n}}, \tag{4}$$

where $x_1 \ldots x_n$ and $y_1 \ldots y_n$ are binary representations of x and y, respectively, and $A_{i,j}$ denote different qubits ($i = 1, 2$ and $j = 1, 2, \ldots N$). We define the N-dimensional DCNOT operator by applying the DCNOT operator on each pair of $A_{1,j}A_{2,j}$:

$$\begin{aligned}|x_1, \ldots, x_n\rangle_{A_{1,1}\ldots A_{1,n}}|y_1, \ldots, y_n\rangle_{A_{2,1}\ldots A_{2,n}} \\ \rightarrow |y_1, \ldots, y_n\rangle_{A_{1,1}\ldots A_{1,n}}|x_1 \oplus y_1, \ldots, x_n \oplus y_n\rangle_{A_{2,1}\ldots A_{2,n}},\end{aligned} \tag{5}$$

that is ($A_1 = A_{1,1}A_{1,2}\ldots A_{1,N}$ and analogously for A_2),

$$|x\rangle_{A_1}|y\rangle_{A_2} \rightarrow |y\rangle_{A_1}|x \oplus y\rangle_{A_2}. \tag{6}$$

Analogously, by applying the inverse DCNOT operator on each pair of $A_{1,j}A_{2,j}$, we have the N-dimensional inverse DCNOT operator:

$$\begin{aligned}|x_1, \ldots, x_n\rangle_{A_{1,1}\ldots A_{1,n}}|y_1, \ldots, y_n\rangle_{A_{2,1}\ldots A_{2,n}} \\ \rightarrow |x_1 \oplus y_1, \ldots, x_n \oplus y_n\rangle_{A_{1,1}\ldots A_{1,n}}|x_1, \ldots, x_n\rangle_{A_{2,1}\ldots A_{2,n}},\end{aligned} \tag{7}$$

that is,

$$|x\rangle_{A_1}|y\rangle_{A_2} \rightarrow |x \oplus y\rangle_{A_1}|x\rangle_{A_2}. \tag{8}$$

2.2. *SDC Protocol 1: Using N-dimensional Bell state*

Protocol 1 uses the following N-dimensional Bell basis,[35] for composite systems of two qu$N$its 1 and 2:

$$|\phi(xy)\rangle_{12} = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{\frac{2\pi i}{N}jx}|j + y\rangle_1|j\rangle_2, \tag{9}$$

where $x, y \in \{0, 1, \ldots, N-1\}$. In the rest of the paper, addition and subtraction inside kets are done modulo N.

The local (acting onto system 1 only) unitary operators $U(xy)$ transform $|\phi(00)\rangle_{12}$ into $|\phi(xy)\rangle_{12}$:

$$(U(xy) \otimes I)|\phi(00)\rangle_{12} = |\phi(xy)\rangle_{12}, \tag{10}$$

where $U(xy) = X^y Z^x$, $X : |j\rangle \rightarrow |j + 1\rangle$ is the shift operator, and $Z : |j\rangle \rightarrow e^{\frac{2\pi i}{N}j}|j\rangle$ is the rotation operator.

The set $\{|\phi(xy)\rangle_{12}\}$ forms an orthonormal basis of completely distinguishable states, and for that reason each of its elements can be used to carry information (xy). The unitary operators $U(xy)$ are then used to encode that information (xy) into the initial state $|\phi(00)\rangle_{12}$.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

In the initialization phase of Protocol 1, Alice, Bob and Charlie share two pairs of entangled qu$N$its $|\phi(00)\rangle_{A_1B}$ and $|\phi(00)\rangle_{A_2C}$, where subscripts A_1 and A_2 denote Alice's two qu$N$its, subscript B denotes Bob's qu$N$it and subscript C denotes Charlie's qu$N$it. The initial quantum state of the composite system is (note that for reasons of simplicity, we drop the subscript A_1BA_2C for Ω-states $|\Omega(0)\rangle$ etc.)

$$|\Omega(0)\rangle = |\phi(00)\rangle_{A_1B} \otimes |\phi(00)\rangle_{A_2C}. \tag{11}$$

Protocol 1 consists of five steps:

(1) *Encoding.* Alice performs unitary operators $U(b_1b_2)$ on qu$N$it A_1 and $U(c_1c_2)$ on qu$N$it A_2 to encode (b_1, b_2) and (c_1, c_2), respectively, like in the original dense coding scheme.[2] After that, the state of the composite system becomes

$$|\Omega(1)\rangle = U_{A_1}(b_1b_2) \otimes U_{A_2}(c_1c_2)|\Omega(0)\rangle = |\phi(b_1b_2)\rangle_{A_1B} \otimes |\phi(c_1c_2)\rangle_{A_2C}. \tag{12}$$

(2) *Locking.* Alice performs the DCNOT operator on qu$N$its A_1A_2 to lock the entanglement channels. The state of the composite system becomes

$$|\Omega(2)\rangle = \mathrm{DCNOT}_{A_1A_2}(|\phi(b_1b_2)\rangle_{A_1B} \otimes |\phi(c_1c_2)\rangle_{A_2C}). \tag{13}$$

(3) *Communication.* Alice sends qu$N$it A_1 to Bob and qu$N$it A_2 to Charlie, like the original dense coding scheme.[2]

(4) *Unlocking.* Bob and Charlie collaborate to perform the inverse DCNOT operator on qu$N$its A_1A_2. The state of the composite system becomes

$$|\Omega(3)\rangle = \mathrm{DCNOT}^{\dagger}_{A_1A_2}|\Omega(2)\rangle = |\phi(b_1b_2)\rangle_{A_1B} \otimes |\phi(c_1c_2)\rangle_{A_2C}. \tag{14}$$

(5) *Decoding.* Bob and Charlie measure qu$N$its A_1B and qu$N$its A_2C in the N-dimensional Bell basis respectively to obtain (b_1, b_2) and (c_1, c_2), like the original dense coding scheme.[2]

The following theorem demonstrates the validity of Protocol 1.

Theorem 1. *Neither Bob nor Charlie alone can learn the encoded information from the states of their subsystems before step 4 (Unlocking) of Protocol 1.*

Proof. After step 2 (*Locking*), the state of the composite system becomes

$$\begin{aligned}|\Omega(2)\rangle = \mathrm{DCNOT}_{A_1A_2}\Bigg(&\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} e^{\frac{2\pi i}{N}jb_1}|j+b_2\rangle_{A_1}|j\rangle_B \\ &\otimes \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} e^{\frac{2\pi i}{N}kc_1}|k+c_2\rangle_{A_2}|k\rangle_C\Bigg) \\ = \frac{1}{N}&\sum_{j,k=0}^{N-1} e^{\frac{2\pi i}{N}(jb_1+kc_1)}|k+c_2\rangle_{A_1}|(j+b_2)\oplus(k+c_2)\rangle_{A_2}|j\rangle_B|k\rangle_C,\end{aligned} \tag{15}$$

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

and the reduced density matrix in subsystem A_1B is

$$\begin{aligned}\rho_{A_1B} &= \sum_{j,k=0}^{N-1} ({}_{A_2}\langle j|_C\langle k|)|\Omega(2)\rangle\langle\Omega(2)|(|j\rangle_{A_2}|k\rangle_C) \\ &= \frac{1}{N^2}\sum_{j,k=0}^{N-1}|j\rangle_{A_1}|k\rangle_{BA_1}\langle j|_B\langle k| \\ &= I_{A_1B}/N^2. \end{aligned} \tag{16}$$

We can calculate the reduced density matrix in subsystem A_2C in the same way and get $\rho_{A_2C} = I_{A_2C}/N^2$. Because ρ_{A_1B} and ρ_{A_2C} are independent of (b_1, b_2) and (c_1, c_2), Bob and Charlie know nothing about the encoded information before step 4 (*Unlocking*). □

2.3. *SDC Protocol 2: Using N-dimensional W state*

Li and Qiu[26] presented a sufficient and necessary condition for a N-dimensional W state to be suitable for perfect teleportation and dense coding, and then they generalized the states of W-class to multi-particle systems with N-dimension:

$$|W(00)\rangle_{123} = \frac{1}{\sqrt{N}}\left(\frac{1}{\sqrt{2}}\sum_{j=1}^{N-1}|j-1\rangle_1(|0j\rangle + |j0\rangle)_{23} + |N-1\rangle_1|00\rangle_{23}\right). \tag{17}$$

Alice uses unitary operators

$$U(xy) = \sum_{j=0}^{N-1} e^{\frac{2\pi i}{N}(j-y)x}|j-y\rangle\langle j| \tag{18}$$

to encode her information:

$$|W(xy)\rangle_{123} = (U(xy)\otimes I\otimes I)|W(00)\rangle_{123}, \tag{19}$$

where $x, y \in \{0, 1, \ldots, N-1\}$.

In the initialization phase of Protocol 2, Alice, Bob and Charlie share two pairs of entangled qu$N$its $|W(00)\rangle_{A_1B}$ and $|W(00)\rangle_{A_2C}$, where subscript A_1 and A_2 denote Alice's two qu$N$its, subscript B denotes Bob's two qu$N$its and subscript C denotes Charlie's two qu$N$its. Protocol 2 consists of five steps:

(1) *Encoding.* Alice performs unitary operators $U(b_1b_2)$ on qu$N$it A_1 and $U(c_1c_2)$ on qu$N$it A_2 to encode (b_1, b_2) and (c_1, c_2), respectively.
(2) *Locking.* Alice performs the DCNOT operator on qu$N$its A_1A_2.
(3) *Communication.* Alice sends qu$N$it A_1 to Bob and qu$N$it A_2 to Charlie.
(4) *Unlocking.* Bob and Charlie collaborate to perform the inverse DCNOT operator on qu$N$its A_1A_2.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

(5) *Decoding.* Bob and Charlie make the von Neumann measurement using the orthogonal states $\{|W(xy)\rangle\}$ on qu$N$its A_1B and qu$N$its A_2C, respectively, to obtain (b_1, b_2) and (c_1, c_2).

The following theorem demonstrates the validity of Protocol 2.

Theorem 2. *Neither Bob nor Charlie alone can learn the encoded information from the states of their subsystems before step 4 (Unlocking) of Protocol 2.*

Proof. The proof is similar to that of Protocol 1. After step 2 (*Locking*), the reduced density matrix

$$\begin{aligned}\rho_{A_1B} = & \frac{1}{N^2}\sum_{j=0}^{N-1}|j\rangle_{A_1}|00\rangle_{BA_1}\langle j|_B\langle 00| \\ & + \frac{1}{2N^2}\sum_{k=1}^{N-1}\sum_{j=0}^{N-1}|j\rangle_{A_1}(|0k\rangle + |k0\rangle)_{BA_1}\langle j|_B(\langle 0k| + \langle k0|) \\ & = \frac{1}{N}\left[\sum_{j=0}^{N-1}|j\rangle_{A_1}\langle j|\right]\otimes\frac{1}{N}\left[|00\rangle_B\langle 00| + \frac{1}{2}\sum_{k=1}^{N-1}(|0k\rangle + |k0\rangle)_B(\langle 0k| + \langle k0|)\right], \quad (20)\end{aligned}$$

and analogously for rho_{A_2C}. Because ρ_{A_1B} and ρ_{A_2C} are independent of (b_1, b_2) and (c_1, c_2), Bob and Charlie know nothing about the encoded information before step 4 (*Unlocking*). □

2.4. *SDC Protocol 3: Using N-dimensional GHZ state*

Protocol 3 uses the following N-dimensional GHZ states[4]:

$$|GHZ(xy)\rangle_{123} = \frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} e^{\frac{2\pi i}{N}jx}|j + y\rangle_1|jj\rangle_{23}, \quad (21)$$

where $x, y \in \{0, 1, \ldots, N-1\}$.

The unitary operators $U(xy)$ transform $|GHZ(00)\rangle_{123}$ into $|GHZ(xy)\rangle_{123}$:

$$(U(xy)\otimes I\otimes I)|GHZ(00)\rangle_{123} = |GHZ(xy)\rangle_{123}, \quad (22)$$

where $U(xy) = X^yZ^x$, $X : |j\rangle \to |j+1\rangle$ is the shift operator, $Z : |j\rangle \to e^{\frac{2\pi i}{N}j}|j\rangle$ is the rotation operator.

Protocol 3 is similar to Protocols 1 and 2, and the proof of its validity is analogous to those of Protocols 1 and 2.

2.5. *Other locking operators*

In this subsection, we introduce another two locking operators: the quantum Fourier transform and the SWAP operator.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

The two-qu$N$it quantum Fourier transform is defined by

$$|x\rangle_{A_1}|y\rangle_{A_2} \to \frac{1}{N}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} e^{\frac{2\pi i}{N^2}(xN+y)(jN+k)}|j\rangle_{A_1}|k\rangle_{A_2}. \tag{23}$$

The SWAP operator simply swaps two qubits (or qu$N$its):

$$|\varphi\rangle_{A_1}|\varphi'\rangle_{A_2} \to |\varphi'\rangle_{A_1}|\varphi\rangle_{A_2}. \tag{24}$$

When the N-dimensional quantum Fourier transform or the SWAP operator are substituted for the DCNOT operator in the above three protocols, the states of Bob's subsystem and Charlie's subsystem after step 2 (*Locking*) are also independent of (b_1, b_2) and (c_1, c_2). Neither Bob nor Charlie alone can learn the encoded information from the states of their subsystems before step 4 (*Unlocking*). Only after performing the inverse quantum Fourier transform or the SWAP operator together, they can achieve (b_1, b_2) and (c_1, c_2) respectively.

When the SWAP operator is used as the locking operator, Protocols 1–3 become simpler. Step 2 (*Locking*) can be omitted. In step 3 (*Communication*), Alice sends qu$N$it A_1 to Charlie and qu$N$it A_2 to Bob. In step 4 (*Unlocking*), Bob and Charlie swap qu$N$it A_1 and qu$N$it A_2. However, if one of the receivers, say Bob, is malicious, he can fool Charlie in step 4 (*Unlocking*) by detaining the qu$N$it that carries (c_1, c_2) (i.e. qu$N$it A_2) and swapping a randomly prepared fake qu$N$it for the qu$N$it that carries (b_1, b_2) (i.e. qu$N$it A_1). Then, Bob would receive (b_1, b_2) from qu$N$its A_1B, but Charlie would not receive achieve (c_1, c_2), because qu$N$it A_2 is still at Bob's side.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

3. Improving the Fairness of SDC

The clients (Bob and Charlie) can decode the messages only by the joint unlocking quantum operation, which requires either a quantum channel, shared entanglement, or direct interaction between them. By "direct interaction", we mean that the clients meet and input their particles into a physical device that can unlock the particles, and then get their respective output particles.

In this section, we discuss the problem of achieving a global unlocking operation between the distant parties (Bob and Charlie) by the means of local operations and classical communication (LOCC), some prior shared resources (such as entanglement, etc.) and/or quantum communication. Also, we discuss the problem of the fairness of the protocol, in case the clients do not trust each other. The solution is a probabilistic protocol based on sequential exchange of information (quantum, or classical with the help of prior shared entanglement) between the clients.

3.1. *The fairness problem*

The problem of achieving a global operation by the actions of spatially distant clients (Bob and Charlie) lies in the fact that the unlocking operation (whether it is the

inverse DCNOT, quantum Fourier transform or SWAP[a]) has a feature that it can entangle initially separable states. Since it is impossible to create entanglement by the means of LOCC only, in order to implement the unlocking operation, Bob and Charlie have to share prior entanglement, or use a quantum channel for one of the clients to send his qu*N*it to the other, who would then perform the unlocking operation on qu*N*its A_1A_2 locally (i.e. at his site).

The problem with any such protocol is that it can be neither simultaneous, nor symmetric, with respect to the two parties involved, which clearly sets an unfair situation.

This is a general problem that arises within asynchronous distributed networks.[36] Namely, in order to achieve the joint operation of the qu*N*its A_1A_2, clients can each perform local measurements, and then send, conditional to the measurement outcomes, (classical or quantum) information to the other. Since the clients are far apart, it is impossible to achieve simultaneous message exchange — the messages are always sent one at a time, from one client to the other (the network is "asynchronous"). This means that, whatever the protocol that achieves the unlocking operation is, there always exists the last message, say from Bob to Charlie. But this means that, prior to sending his *last message* to Charlie, Bob has all the right (quantum) information needed to obtain his message (b_1, b_2), while Charlie does not. This is obviously unfair, with respect to Charlie.

3.2. *The solution*

The situation is similar to the one presented in the contract signing problem.[34] The solution, similar to the one proposed for quantum contract signing,[37] is to perform the unlocking operation on qu*N*its A_1 and A_2 (i.e. arrays of qubits $A_{1,1} \ldots A_{1,n}$ and $A_{2,1} \ldots A_{2,n}$) in series of steps, such that after each step, one pair of qubits are unlocked. In the course of the unlocking stage, the clients increase their probabilities to obtain the needed classical information ((b_1, b_2) and (c_1, c_2), respectively), such that at each step one client has slightly higher probability of successful recovery than the other. Therefore, the protocol is probabilistic, and also fair in the sense that at each step, one client is only slightly privileged over the other.

The information can be transferred in two equivalent ways: either by sending quantum information (qubits) via a quantum channel, or by teleporting qubits' quantum states using LOCC and shared entanglement. Without the loss of generality, we will assume that the clients are exchanging the actual qubits, rather than teleporting their states. The only difference between the two cases is that in the latter

[a]The case of the SWAP operation is particularly interesting. Namely, after Bob and Charlie swap the states of qubits A_1 and A_2, the very qubits do not become entangled with each other. Nevertheless, the distant sites of Bob and Charlie do become entangled: qubits A_1 are now entangled with qubits C, while qubits A_2 are entangled with qubits B. The other way to look at this is through the no-cloning theorem: there is no way for Bob to learn the unknown state of the system A_1B and transfer (swap) its partial state of A_1 together with its entanglement with B using only LOCC, without prior entanglement shared with Charlie.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

case, the clients use the previously shared entanglement and exchange classical instead of quantum information.

The main problem in this approach is that a client, say Charlie, cannot be sure if Bob sent him the right information or not. Therefore, we introduce additional $2s$ "*control qubits*" that are used by the clients to check each other's honesty during the unlocking stage: s control qubits would be joined with n "*message qubits*" that carry the message (b_1, b_2), to form $(n+s)$ qubits $A_{1,1} \ldots A_{1,n+s}$, and given to Bob; other s control qubits joined with n message qubits that carry (c_1, c_2), to form $A_{2,1} \ldots A_{2,n+s}$ and given to Charlie.

The position $C\text{pos}_1(j) \in \{1, 2, \ldots, n+s\}$ of each control qubit $j = 1, 2, \ldots, s$ within the $(n+s)$ qubits $A_{1,1} \ldots A_{1,n+s}$ is chosen randomly by Alice and this information (s integers $C\text{pos}_1(j)$) is given to Bob; analogously, random positions $C\text{pos}_2(j)$ of control qubits from $A_{2,1} \ldots A_{2,n+s}$ are given to Charlie.

The control qubits are prepared by Alice in pure states and are uncorrelated from the message qubits. The state of each control qubit is randomly chosen by Alice from a publicly known set $\mathcal{S}$ of pure states. $\mathcal{S}$ is the union of two mutually unbiased bases, computational Z basis $\{|0\rangle, |1\rangle\}$ and rotated X basis $\{|+\rangle, |-\rangle\}$, with $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$; therefore, $\mathcal{S} = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. The pure state of jth control qubit in the position $C\text{pos}_1(j)$ among qubits $A_{1,1} \ldots A_{1,n+s}$ is $C\text{state}_1(j)$, with $C\text{state}_1(j) \in \mathcal{S}$. Analogously, the states of control qubits from $A_{2,1} \ldots A_{2,n+s}$ are encoded by $C\text{state}_2(j) \in \mathcal{S}$. Alice gives the information $\{C\text{state}_1(j) | j = 1, 2, \ldots, s\}$ about the states of control qubits from $A_{1,1} \ldots A_{1,n+s}$ to Bob, and the information $\{C\text{state}_2(j) | j = 1, 2, \ldots, s\}$ about the states of control qubits from $A_{2,1} \ldots A_{2,n+s}$ to Charlie.

Therefore, prior to Alice sending the message to clients (via sending the qubits $A_{1,1} \ldots A_{1,n+s}$ to Bob and $A_{2,1} \ldots A_{2,n+s}$ to Charlie), Bob has, apart from qu$N$it B, the information $\{C\text{pos}_1(j) | j = 1, 2, \ldots, s\}$ of the positions and $\{C\text{state}_1(j) | j = 1, 2, \ldots, s\}$ of the states of the control qubits from $A_{1,1} \ldots A_{1,n+s}$, and analogously for Charlie.

During the initialization phase, Alice, Bob and Charlie share two pairs of entangled qu$N$its, denoted as

$$|\phi(00)\rangle_{A_1 B} \otimes |\phi(00)\rangle_{A_2 C}. \tag{25}$$

Alice and Bob also share classical information of the control qubits: $\{C\text{pos}_1(j) | j = 1, 2, \ldots, s\}$ and $\{C\text{state}_1(j) | j = 1, 2, \ldots, s\}$. Alice and Charlie also share classical information $\{C\text{pos}_2(j) | j = 1, 2, \ldots, s\}$ and $\{C\text{state}_2(j) | j = 1, 2, \ldots, s\}$.

The protocol of SDC of classical messages (b_1, b_2) and (c_1, c_2) works as follows:

(1) *Encoding.* Alice performs unitary operators $U(b_1 b_2)$ on qu$N$it A_1 and $U(c_1 c_2)$ on qu$N$it A_2 to encode (b_1, b_2) and (c_1, c_2), respectively.
(2) *Locking.* Alice joins the s control qubits with n message qubits $A_{1,1} \ldots A_{1,n}$, to form the ordered set of $(n+s)$ qubits $A_{1,1} \ldots A_{1,n+s}$. The jth control qubit is

prepared in $C\text{state}_1(j) \in \mathcal{S}$ and is in the position $C\text{pos}_1(j)$, while the relative positions of the n message qubits are the same as before. Analogously, she forms the set $A_{2,1} \ldots A_{2,n+s}$.

For $j = 1$ to $n + s$, Alice applies the locking operator on $A_{1,j}A_{2,j}$.

(3) *Communication.* Alice sends $(n + s)$ qubits $A_{1,1} \ldots A_{1,n+s}$ to Bob, and $(n + s)$ qubits $A_{2,1} \ldots A_{2,n+s}$ to Charlie.

(4) *Unlocking.* For $j = 1$ to $n + s$, Bob sends qubit $A_{1,j}$ to Charlie, and then Charlie returns $A_{1,j}$ to Bob after performing the unlocking operator on $A_{1,j}A_{2,j}$ at his site.

If $\exists k, C\text{pos}_1(k) = j$ (i.e. if the jth qubit given to Bob is a controlled one), Bob measures $A_{1,j}$ in either X or Z basis, according to $C\text{state}_1(k)$. If the measurement result does not match $C\text{state}_1(k)$, he knows that Charlie did not return the real $A_{1,j}$ and stops the unlocking stage.

Analogously, if $\exists k, C\text{pos}_2(k) = j$, Charlie measures $A_{2,j}$ in either X or Z basis, according to $C\text{state}_2(k)$. If the measurement result does not match $C\text{state}_2(k)$, he knows that Bob did not send the real $A_{1,j}$ and stops the unlocking stage.

The remaining n qubits received at Bob's site form the ordered set $A_{1,1} \ldots A_{1,n}$, and the remaining n qubits received at Charlie's site form the ordered set $A_{2,1} \ldots A_{2,n}$.

(5) *Decoding.* Bob and Charlie measure qu$N$its A_1B and qu$N$its A_2C in the Bell basis respectively to achieve (b_1, b_2) and (c_1, c_2).

The unlocking of qubits $A_{1,1} \ldots A_{1,n+s}$ and $A_{2,1} \ldots A_{2,n+s}$ is done in $(n + s)$ steps, such that in each step one pair of qubits is unlocked. The order of qubits must be maintained. Without the loss of generality, we assume that the unlocking operation is done at Charlie's site. We will assume that Bob is an honest client who sends qubits $A_{1,1} \ldots A_{1,n+s}$ to Charlie, as long as he is convinced that Charlie is returning the exact resulting qubit after the local unlocking operation at Charlie's site. The way to check if Charlie is indeed doing so is the following: in each step $C\text{pos}_1(j)$ of the unlocking stage (i.e. when the qubit given by Alice to Bob is a controlled one), Bob measures the state of the qubit received back from Charlie during that step in one of the two mutually unbiased bases — Z if $C\text{state}_1(j)$ is in $\{|0\rangle, |1\rangle\}$, X otherwise. If the measurement result matches the classical information $C\text{state}_1(j)$, i.e. if Bob's outcome is consistent with Charlie returning the control qubit in $C\text{state}_1(j)$, Bob continues with the unlocking stage. Otherwise, it means that Charlie did not return the control qubit in $C\text{state}_1(j)$ and Bob stops the unlocking stage. Since Charlie does not know the positions of control qubits from $A_{1,1} \ldots A_{1,n+s}$, he has to return all of the resulting qubits to Bob. Otherwise, he will inevitably return some qubits as control ones in states different from those prepared by Alice, and Bob will, with high probability, be able to detect it.

Note that the whole analysis is done for the ideal case where no measurement errors or decoherence effects occur. The existence of measurement errors will set the threshold value $\eta > 0$ for the allowed number of wrong results obtained on control

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

qubits (results inconsistent with Charlie sending the control qubits in $C\text{state}_1(j)$), which will increase the number s of the control qubits. After receiving, in the course of exchange, $k \leq s$ controlled qubits, we require that not more than ηk wrong results are obtained. The parameter η is determined by the experimental set-up, which sets the probability of obtaining the wrong result when the right qubit is sent.

The protocol is *optimistic*[38]: if both clients are honest, if they execute the protocol by unlocking the qubits Alice gave them, upon finishing the unlocking stage both parties will have the whole information sent by Alice.

Unfortunately, if we insist on the *perfect* fidelity of data transmission, situations when one client (say Charlie) knows that he obtained the whole information, while the other has not, clearly puts Bob in a disadvantaged situation. For example, if the unlocking operation is done at Charlie's site, Charlie can decide not to return the last resulting qubit to Bob, which would leave him without the whole state $|\phi(b_1 b_2)\rangle_{A_1 B}$, and thus (b_1, b_2), in case $C\text{pos}_1(s) < n + s$. The perfect fidelity was for the same reasons relaxed for quantum contract signing as well, by introducing the factor $\alpha < 1$ of the required fraction of correct results for the total number N of qubits sent to a client.[37] Therefore, we will also introduce factor $\alpha < 1$ and require that a client has correct values of $\alpha(2n)$ message bits sent to him by Alice.

The above protocol is clearly *probabilistic*. At each step of the unlocking stage, Bob has a finite probability, which increases with the execution of the protocol, to obtain correct values of $2\alpha n$ bits of the message (b_1, b_2) that Alice sent him, and analogously for Charlie.

Due to its probabilistic nature, during the execution of the protocol one client is always privileged over the other. By privileged, we mean that one client has higher probability of obtaining the required fraction α of the message Alice encoded for him, than the other. The difference between the clients' probabilities to obtain the information sent by Alice is due to the random distribution of control qubits. On the other side, since neither of the clients know the distribution of both sets of control qubits, even though one of them might be privileged over the other at a certain step of the unlocking stage, he would not know that: the protocol is *a priori* symmetric with respect to the clients.

Yet, the protocol is *fair*[39]: throughout the execution, one client is only slightly privileged over the other, the difference being smaller with the increase of s and could be made arbitrarily small for big enough s. The argument here is exactly the same as the one presented for the fairness of the quantum contract signing protocol.[37]

Namely, the role of the control qubits is to signal possible *cheating* of a client. By cheating, we mean not sending the qubits received by Alice. For example, if one of the clients, say Charlie, starts sending qubits each randomly in one of four states $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$, he will send wrong message qubits A_1 sent by Alice, but also wrong control qubits. Therefore, for each control qubit he will have a finite probability of 1/2 of being detected cheating (he did not send the right control qubit). Therefore, since Bob's probability p_w to detect a wrong qubit (Charlie's cheating) is

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

approaching to one exponentially fast, $p_w = 1 - (1/2)^m$, where m is a number of controlled qubits that are sent as random, it is not difficult to estimate the expected difference between Bob and Charlie, depending on the total number n of the message qubits to be sent from one client to another and the number s of the controlled qubits given to each client.

One simply has to estimate the expected value for m, the expected number $\langle m \rangle$ of wrongly sent control qubits after which Bob notices the cheating and stops communication. During that time, Bob received certain number w of random qubits as the message qubits, while Charlie was still receiving the proper ones. Therefore, he is in advantage by having about $w/2$ more correct values for message bits (even when guessing bit values at random, one has $1/2$ of the probability to be correct). The value w is a simple function of m, $w = w(m)$, and the dependence on m can be straightforwardly determined by n and s. For big enough n, $w(\langle m \rangle) \ll n$ and therefore Charlie's advantage is negligible.

Of course, Charlie can try to send qubits in some other states (completely random, etc.), but as long as he is not sending the right qubits in the right states, there will be a finite probability $p > 0$ that Bob detects cheating in a single-shot measurement, so that his probability to detect cheating will again exponentially fast approach to one in the number of wrongly sent control qubits, making Charlie's privilege arbitrarily small, for big enough n. Charlie can try to decrease the value of p, but can never make it zero. Otherwise, he would be able to perfectly distinguish between the non-orthogonal states and this would violate the security of the BB84 cryptographic protocol,[40] for example.

In the case of quantum contract signing[37] the role of message and control qubits was given to the same N qubits sent by a trusted party (in our case Alice) to clients (in our case Bob and Charlie), while in the case of SDC the two roles are given to separate sets of qubits. This will introduce a slight change in the expressions for the probabilities involved in calculation, but this change is minor, conceptually straightforward to calculate and does not affect the main result of protocol's fairness. Nevertheless, knowing the exact expressions for the probabilities in case of different cheating strategies is of crucial importance, and may technically be quite non-trivial, which clearly presents interesting and challenging topics for future research.

The fairness condition can be even straightened, as was done in Ref. 37, by requiring the negligible *probability to cheat*: the probability that one client has $2\alpha n$ right bits, while the other does not. This quantity may be quite relevant in various possible scenarios, like in the case of signing contracts for buying and selling the goods on the market (see Ref. 37). While for a fixed value of α the probability to cheat may be as high as $1/4$, if it is unknown to the clients and chosen randomly by Alice from a certain interval $I_\alpha \subset (1/2, 1)$, the probability to cheat can be made as small as needed, for big enough n (see Ref. 37).

The locking and unlocking operators performed on two qu$N$its (i.e. two arrays of qubits) in this section are actually products of two-qubit locking and unlocking

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

operators. The N-dimensional DCNOT operator and the SWAP operator are of this kind. But the N-dimensional quantum Fourier transform cannot be done qubit(pair)-by-qubit(pair). Therefore, the quantum Fourier transform cannot be used to implement the SDC protocol proposed in this section.

4. Improving the Security of SDC

We now consider the security of SDC. We assume that in the initialization phase, the entangled pairs have been securely distributed among Alice, Bob and Charlie. Because only part of the entangled pairs travel through the quantum channel, an outsider knows nothing about the encoded information. If the two receivers are honest and follow the protocols exactly, they must collaborate to achieve their respective information. However, if one of the receivers, say Bob, is dishonest and has the ability to intercept and resend the qubits going through the quantum channel between Alice and Charlie, he can obtain (b_1, b_2) without collaborating with Charlie by the following intercept–resend attack: (1) intercept qu$N$it A_2; (2) perform the unlocking operation on qu$N$its A_1A_2; (3) measure qu$N$its A_1B to obtain (b_1, b_2); (4) perform the locking operation on qu$N$its A_1A_2; (5) send qu$N$it A_2 back to Charlie.

To detect such a cheating behavior, we insert additional $2r$ "*detect qubits*" into the array of $2n$ message qubits during the communication phase: r detect qubits would be joined with n message qubits that carry the message (b_1, b_2), to form $(n + r)$ qubits $A_{1,1} \ldots A_{1,n+r}$, and given to Bob; other r detect qubits joined with n message qubits that carry (c_1, c_2), to form $A_{2,1} \ldots A_{2,n+r}$ and given to Charlie.

The position $D\text{pos}_1(j) \in \{1, 2, \ldots, n + r\}$ of each detect qubit $j = 1, 2, \ldots, r$ within the $(n + r)$ qubits $A_{1,1} \ldots A_{1,n+r}$ is chosen randomly by Alice, and analogously for $D\text{pos}_2(j)$ of each detect qubit within $A_{2,1} \ldots A_{2,n+r}$.

The state of each detect qubit is randomly chosen by Alice from set $\mathcal{S} = Z \cup X = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. The pure state of jth detect qubit in the position $D\text{pos}_1(j)$ among qubits $A_{1,1} \ldots A_{1,n+r}$ is $D\text{state}_1(j)$, with $D\text{state}_1(j) \in \mathcal{S}$. Analogously, the states of detect qubits from $A_{2,1} \ldots A_{2,n+r}$ are encoded by $D\text{state}_2(j) \in \mathcal{S}$.

Only after the transmission of all the message qubits and detect qubits, Alice tells Bob and Charlie the positions and the bases of the detect qubits and requires them to return the results of the measurements performed on the detect qubits. Therefore, Alice can check if after the transmission the states of the detect qubits have been altered by a dishonest client.

Because the position of the detect qubits in the array are chosen randomly by Alice, curious Bob does not know which qubits are the message qubits. If his intercept–resend attack involves a detect qubit, the state of the detect qubit may probably be changed, the probability of being detected grows exponentially with the increase of r and could be made arbitrarily large for big enough r.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

Table 1. Comparison between different protocols.

Protocol	Properties	Extra cost
protocol in Sec. 2 and Ref. 33		
protocol in Sec. 3	fair	s control qubits
protocol in Sec. 4	secure	r detect qubits
protocol in Appendix	fair, secure	s control qubits + r detect qubits

During the initialization phase, Alice, Bob and Charlie share two pairs of entangled qu$N$its, denoted as

$$|\phi(00)\rangle_{A_1B} \otimes |\phi(00)\rangle_{A_2C}. \tag{26}$$

The protocol of SDC of classical messages (b_1, b_2) and (c_1, c_2) works similarly as in the previous case. For completeness, in Appendix A we present a detailed description of the combination of the two strategies that assure fair and secure SDC protocol.

We summarize the above protocols in Table 1 in order to demonstrate their qualitative and quantitative differences. The original SDC protocol in Ref. 33 and the N-dimensional SDC protocol in Sec. 2 are vulnerable to intercept–resend attacks and require the receivers to perform the unlocking operation at the same site, otherwise the fairness problem arises. The improved SDC protocol in Sec. 3 can guarantee fairness in the unlocking phase even if the receivers are separated spatially, by implementing the unlocking operation in series of steps of communication and introducing additional control qubits for cheat detection. On the other hand, the improved SDC protocol in Sec. 4 can guarantee security against the intercept–resend attack in the communication phase, by introducing additional detect qubits in the communication phase. The protocol in Appendix incorporates both fairness and security properties. We can also see from the table that the cost of the protocol increase with these improvements.

5. Applications to Contract Signing

Contract signing[34] is a security task involving two parties, Charlie and Bob, that do not trust each other and want to exchange a common contract signed with each other's signature. At the end of the protocol Charlie should have the contract signed by Bob and vice versa. The purpose of the signed contract is to bind the parties to the terms of the contract, which can be enforced by a judge (Alice). The real challenge to this problem is when both Charlie and Bob are physically apart and want to remotely sign the contract. This situation is becoming more and more common due to e-business, and may lead to fraud. For instance, if Bob gets the contract signed by Charlie without committing himself, Charlie and Bob are in an unfair situation. By having Charlie's commitment, Bob is able to appeal to Alice to enforce the contract, while Charlie has no means to do the same, since he does not possess the contract signed by Bob. Note that even if Bob did not commit, but having Charlie's

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

commitment, puts him in a position to later in time choose whether to bind the contract or not, while Charlie has no power to do either of the two.

A simple solution to this unfair situation is to have a trusted third party (again Alice) mediating the transaction — Bob sends to Alice the contract signed by him and Charlie does the same; then Alice exchanges the contracts only after she has received both of the commitments. Note that this procedure increases significantly the cost of remote contract signing, as Alice's time and resources are expensive. What is particularly costly is Alice being constantly online and alert waiting for the clients to contact her. Also, avoiding the communication with the trusted party at the very moment of determining a contract and committing to it removes the danger of overloading Alice and creating a bottleneck.

Unfortunately, it has been shown that the attendance of Alice is mandatory,[34,36] if the protocol is to fulfil the following two important properties:

- *fairness*: either both parties get each other' commitment or none gets;
- *viability*: if both parties behave honestly, they will both get each other's commitments.

One way to overcome this difficulty is to consider *optimistic protocols* that do not require communication with Alice unless something wrong comes up.[38] Another workaround is to relax the fairness condition, allowing one agent to have ϵ more probability of binding the contract than the other agent (*probabilistic fairness*). In this case, for an arbitrary small ϵ solutions have been found where the number of exchanged messages between the agents is minimized.[39]

In this section, we present a probabilistically fair quantum protocol based on a secure SDC for remote agents, such as the one described in Appendix A (a combination of protocols presented in Secs. 3.2 and 4). In such protocol, Alice, the trusted party, does not interact with the signing parties while they (Bob and Charlie) are determining the contract and then committing to it (the Exchange Phase). Note that in the below construction of a contract signing scheme, a SDC protocol is used iteratively many times as a *black box*. Thus, *any* protocol that achieves simultaneous transmission of (densly) coded information that satisfies the fairness and the security conditions, as presented in Secs. 3 and 4, respectively, allows for constructing probabilistically fair contract signing scheme.

In order to lower the use of resources, it is possible to map long contracts into messages of a small fixed size, say of k bits. Such short messages (digests) are obtained by so called hash functions (such as SHA1) and are well established in the field of cryptography.[41] Hash functions are not injective, that is, there exist pairs of different messages x, x' with the same digest d. Nevertheless, given a message x and its digest d, it is computationally hard to find a different message $x' \neq x$ with the same digest d. For this reason, digests can be used to identify a message. From this point on, instead of contracts themselves we consider their digests, obtained by some hash function, with k bits, say $(b_1 \dots b_k)$, where $b_i \in \{0, 1\}$.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

We also assume that Alice can sign messages using some public key signing scheme (such as DSS, for more detail see Ref. 41). In short, a public key signature scheme for Alice is a pair of functions (sig, ver) together with a pair of keys $(A, \hat{A})$ where A is the private key of Alice (known only by Alice) and $\hat{A}$ is the corresponding public key (known by Alice, Bob and Charlie). For Alice to sign a message m, she uses her private key A and obtains the signature $sig_A(m)$. Bob (or Charlie) verifies if $sig_A(m)$ is the message m signed by Alice by checking whether $ver_{\hat{A}}(m, sig_A(m)) = 1$. If $ver_{\hat{A}}(m, sig_A(m)) \neq 1$ then $sig_A(m)$ does not correspond to the signature of Alice over m. Signing can be seen as an encryption, unique to Alice: only she can do it using her private key A. But Bob and Charlie can, given the message m and a signature s, by using Alice's public key $\hat{A}$ verify whether s is indeed the signature of m by Alice or not.

In our contract signing protocol, Alice does not know *a priori* to which particular contract/digest $(b_1 \ldots b_k)$ Bob and Charlie are going to agree upon. However, at the end of the protocol Alice (and also Bob and Charlie) needs some irrefutable proof of the particular contract that was agreed upon. Moreover, we do not want Alice to be contacted during the exchange phase. A solution for this problem is for Alice to produce $4k$ triples $\{(b, i, \text{Bob}), (b, i, \text{Charlie}) : b \in \{0, 1\}, i = 1 \ldots k\}$ that can be used to represent any particular contract $(b_1 \ldots b_k)$ that Bob and Charlie will agree upon latter. Next, Alice signs those $4k$ triples and prepares $2k$ SDC protocols for Bob and Charlie, where in each of these SDC protocols the messages to be simultaneously received are $sig_A(b, i, \text{Charlie})$ by Bob and $sig_A(b, i, \text{Bob})$ by Charlie, with $b \in \{0, 1\}$ and $i \in \{1 \ldots k\}$.

Thus, Bob can enforce the contract $(b_1 \ldots b_k)$ against Charlie (and no other person) if he shows to Alice the signatures $sig_A(b_1, 1, \text{Charlie}) \ldots sig_A(b_k, k, \text{Charlie})$. By using SDC, it is possible to force Bob to obtain this information (and no other) only with the collaboration of Charlie, and vice versa. In detail the contract signing protocol based on SDC works as follows:
Initialization Phase:

(1) Alice signs the following $4k$ messages: $(0, i, \text{Bob})$, $(1, i, \text{Bob})$, $(0, i, \text{Charlie})$, $(1, i, \text{Charlie})$ with $i = 1, 2, \ldots, k$.
(2) Alice arranges for $2k$ different SDC's for the case of distant parties such that the message to be sent to Bob in one of such SDC's is $sig_A(b, i, \text{Charlie})$ and to Charlie is $sig_A(b, i, \text{Bob})$ for $i = 1, 2, \ldots, k$ and $b = 0, 1$.

Exchange Phase:

(1) Bob and Charlie agree on contract $(b_1 \ldots b_k) \in \{0, 1\}^k$.
(2) For $i = 1$ to k, Bob and Charlie collaborate to obtain from the entangled qu$N$its of the $2k$ SDC's the messages $sig_A(b_i, i, \text{Charlie})$ and $sig_A(b_i, i, \text{Bob})$, respectively, and ignore the remaining quantum data sent by Alice. Thus, at the end of the SDC's Bob has $sig_A(b_i, i, \text{Charlie})$, for $i = 1 \ldots k$, and mutatis mutandis for Charlie.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

Binding Phase:

(1) Alice enforces contract $(b_1 \dots b_k)$ when either Bob presents $sig_A(b_1, 1, \text{Charlie}) \dots sig_A(b_k, k, \text{Charlie})$ or Charlie presents $sig_A(b_1, 1, \text{Bob}) \dots sig_A(b_k, k, \text{Bob})$.

The size of each signature, say $sig_A(b_i, i, \text{Charlie})$, is given by the length of a message a client receives in each SDC, which is $2n$ bits (the factor 2 comes from the fact that the coding is dense). Since each SDC is probabilistic, a client only needs to present to Alice $\alpha_i(2n)$ bits, with $\frac{1}{2} < \alpha_i < 1$, of the signature $sig_A(b_i, i, \text{Charlie})$ for each $i = 1 \dots k$. Moreover, if each α_i is random we achieve even stronger fairness condition, namely the *expected* probability to cheat on each SDC can be made arbitrarily small (the probability to cheat is the probability that an agent obtains at least $\alpha_i(2n)$ bits of a signature and the other does not[37]).

The contract signing protocol described above, unlike the one presented in Ref. 37, allows for Bob and Charlie to determine the contract after the initialization phase. This is due to the fact that public key signatures were introduced. However, this introduction leads to a poorer security assumption, as public key signatures are only computationally secure (as well as hash functions).

We can improve the security of the above protocol by removing both hash functions and public key signatures. Removing hash function accounts to consider a large enough k such that all potential contracts by Bob and Charlie would fit in k bits. To remove public key signatures, which are not perfectly secure, Alice can, during the initialization phase, share a symmetric key k_{AB} with Bob and another, k_{AC}, with Charlie. These symmetric keys might be made perfectly secure by using one-time pad cryptosystems (see, for example[41]). Then, in each SDC the message that Bob receives is $k_{AB}(b_i, i, r_i(b_i))$ and the message that Charlie receives is $k_{AC}(b_i, i, r_i(b_i))$. Here, $k_{AB}(m)$ is the encryption of m with the symmetric key k_{AB} and $r_i(b_i)$ is a random string for each i and b_i, sampled and known only by Alice, that is associated to a contract between Bob and Charlie. So, for Bob to enforce contract $(b_1 \dots b_k)$ against Charlie, he has to present to Alice the random numbers $r_1(b_1) \dots r_k(b_k)$. Note that in this case Alice has to store all these random numbers $r_i(b_i)$ in her private memory keeping in mind to whom they are associated. In this way the protocol's perfect security is obtained by the laws of physics, which is stronger than computational security used in classical protocols.

6. *N*-Dimensional Simultaneous Teleportation

Dense coding and quantum teleportation are dual protocols intimately linked to each other, both in their purpose as well as in the construction. The former is used to transmit (densely coded) classical information, while the latter transfers quantum states (i.e. quantum information). In introduction, as a motivation for SDC the following situation was described: an agent, say a boss of a company, needs for two employees to simultaneously carry out two different confidential tasks, such that

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

each employee is unaware of the other's activity. SDC allows that the two tasks, encoded by classical information, are confidentially communicated to the employees. Similarly, one could imagine a situation in which the employees are requested to each execute a predetermined quantum protocol (say, a quantum computation algorithm) for a given confidential initial state. The solution for such situation is achieved by a simultaneous teleportation protocol.

Simultaneous teleportation was proposed by Wang *et al.*,[32] in which all the receivers simultaneously obtain their respective quantum states from the sender. In their scheme, the sender first performs a locking operation to entangle the particles from two independent quantum entanglement channels, and therefore the receivers cannot restore their quantum states separately before performing the unlocking operation together. The locking operator is composed of the Hadamard and the CNOT operators.

Reference 33 showed that the quantum Fourier transform can alternatively be used as the locking operator in simultaneous teleportation. In this section, we further investigate simultaneous teleportation of qu$N$its using the N-dimensional quantum Fourier transform.

In the task of N-dimensional simultaneous teleportation, Alice intends to teleport the unknown qu$N$its $|\varphi_t\rangle_{T_t} = \sum_{s=0}^{N-1} \alpha_{t,s}|s\rangle_{T_t}$ to Bob_t $(1 \le t \le M)$ under the condition that all the receivers must collaborate to simultaneously obtain $|\varphi_t\rangle_{T_t}$.

In the initialization phase, Alice shares with each Bob_t a pair of entangled qu$N$its $|\phi(00)\rangle_{A_t B_t}$, where subscript A_t denotes Alice's qu$N$it, and subscript B_t denotes Bob_t's qu$N$it. The initial quantum state of the composite system is

$$\begin{aligned} |\chi(0)\rangle &= \bigotimes_{t=1}^{M} |\phi(00)\rangle_{A_t B_t} \bigotimes_{t=1}^{M} |\varphi_t\rangle_{T_t} \\ &= \frac{1}{\sqrt{N^M}} \sum_{j=0}^{N^M-1} |j\rangle_{A_1 \ldots A_M} |j\rangle_{B_1 \ldots B_M} \bigotimes_{t=1}^{M} |\varphi_t\rangle_{T_t}, \end{aligned} \tag{27}$$

where j is a base-N-number which can be written as $j_1 j_2 \ldots j_M, j_t \in \{0, 1, \ldots, N-1\}$.

The protocol for simultaneous teleportation of qu$N$its consists of five steps:

(1) *Locking.* Alice performs the N-dimensional quantum Fourier transform

$$|j\rangle_{A_1 \ldots A_M} \to \frac{1}{\sqrt{N^M}} \sum_{k=0}^{N^M-1} e^{\frac{2\pi i}{N^M} jk} |k\rangle_{A_1 \ldots A_M}, \tag{28}$$

on qu$N$its $A_1 \ldots A_M$ to lock the entanglement channels. After that, the state of the composite system becomes

$$|\chi(1)\rangle = \frac{1}{\sqrt{N^M}} \sum_{j=0}^{N^M-1} QFT_{A_1 \ldots A_M} |j\rangle_{A_1 \ldots A_M} |j\rangle_{B_1 \ldots B_M} \bigotimes_{t=1}^{M} |\varphi_t\rangle_{T_t}$$

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

$$= \frac{1}{\sqrt{N^M}} \sum_{j=0}^{N^M-1} \frac{1}{\sqrt{N^M}} \sum_{k=0}^{N^M-1} e^{\frac{2\pi i}{N^M}jk} |k\rangle_{A_1 \ldots A_M} |j\rangle_{B_1 \ldots B_M} \bigotimes_{t=1}^{M} |\varphi_t\rangle_{T_t}. \tag{29}$$

(2) *Measuring.* Alice measures each pair of qu$N$its A_tT_t in the N-dimensional Bell basis.

$$\begin{aligned}
&\bigotimes_{t=1}^{M} {}_{A_tT_t}\langle\phi(x_t y_t)|\chi(1)\rangle \\
&= \frac{1}{N^M} \left(\bigotimes_{t=1}^{M} \sum_{j=0}^{N-1} e^{-\frac{2\pi i}{N}jx_t} {}_{A_t}\langle j + y_t|_{T_t}\langle j| \right) \\
&\quad \times \left[\sum_{k=0}^{N^M-1} \left(\bigotimes_{t=1}^{M} |k_t\rangle_{A_t} |\varphi_t\rangle_{T_t} \right) \frac{1}{\sqrt{N^M}} \sum_{j=0}^{N^M-1} e^{\frac{2\pi i}{N^M}jk} |j\rangle_{B_1 \ldots B_M} \right] \\
&= \frac{1}{N^M} \sum_{k=0}^{N^M-1} \left(\prod_{t=1}^{M} \sum_{j=0}^{N-1} e^{-\frac{2\pi i}{N}jx_t} \langle j + y_t|k_t\rangle\langle j| \sum_{s=0}^{N-1} \alpha_{t,s}|s\rangle \right) QFT_{B_1 \ldots B_M} |k\rangle_{B_1 \ldots B_M} \\
&= \frac{1}{N^M} QFT_{B_1 \ldots B_M} \sum_{k=0}^{N^M-1} \bigotimes_{t=1}^{M} e^{-\frac{2\pi i}{N}x_t(k_t - y_t)} \alpha_{t,k_t - y_t} |k_t\rangle_{B_t} \\
&= \frac{1}{N^M} QFT_{B_1 \ldots B_M} \sum_{k=0}^{N^M-1} \bigotimes_{t=1}^{M} e^{-\frac{2\pi i}{N}x_t k_t} \alpha_{t,k_t} |k_t + y_t\rangle_{B_t} \\
&= \frac{1}{N^M} QFT_{B_1 \ldots B_M} \sum_{k=0}^{N^M-1} \bigotimes_{t=1}^{M} \alpha_{t,k_t} X^{y_t} (Z^\dagger)^{x_t} |k_t\rangle_{B_t} \\
&= \frac{1}{N^M} QFT_{B_1 \ldots B_M} \bigotimes_{t=1}^{M} X^{y_t} (Z^\dagger)^{x_t} \sum_{s=0}^{N-1} \alpha_{t,s}|s\rangle_{B_t},
\end{aligned} \tag{30}$$

where $X : |j\rangle \to |j+1\rangle$ is the shift operator, $Z : |j\rangle \to e^{\frac{2\pi i}{N}j}|j\rangle$ is the rotation operator.

If the measurement result of qu$N$its A_tT_t is $|\phi(x_t y_t)\rangle$, the state of qu$N$its $B_1 \ldots B_M$ collapses into

$$|\chi(2)\rangle = QFT_{B_1 \ldots B_M} \bigotimes_{t=1}^{M} X^{y_t} (Z^\dagger)^{x_t} |\varphi_t\rangle_{B_t}. \tag{31}$$

(3) *Communication.* Alice sends the measurement result (x_t, y_t) to each Bob$_t$.

(4) *Unlocking.* All the receivers collaborate to perform the inverse quantum Fourier transform on qu$N$its $B_1 \ldots B_M$, and the state of qu$N$its $B_1 \ldots B_M$ becomes

$$|\chi(3)\rangle = QFT^\dagger_{B_1 \ldots B_M} |\chi(2)\rangle = \bigotimes_{t=1}^{M} X^{y_t} (Z^\dagger)^{x_t} |\varphi_t\rangle_{B_t}. \tag{32}$$

(5) *Recovering.* Each Bob$_t$ performs $Z^{x_t}(X^\dagger)^{y_t}$ on qu$N$it B_t to obtain $|\varphi_t\rangle$.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

7. Conclusions

Dense coding[2] and teleportation[35] are important quantum communication tasks. Simultaneous dense coding[33] and simultaneous teleportation,[32] which guarantee that the receivers simultaneously achieve their respective information from one sender, may be relevant and useful for improvement of some models or tasks of quantum communication. In this paper, we have given a number of new results on SDC and teleportation. More specifically, we have given three protocols for SDC utilizing different N-dimensional quantum states (i.e. Bell state, W state, and GHZ state). Besides the quantum Fourier transform, we have introduced two new locking operators (i.e. the double controlled-NOT operator and the SWAP operator) for SDC. Then we have analyzed the fairness and the security of the SDC protocol and proposed a protocol which guarantees both the fairness and the security, thus allowing for mutually distant receivers to execute the protocol. We have shown that any fair SDC scheme that is secure against intercept–resend attack can be used to implement a fair contract signing protocol. In addition, we have shown that the N-dimensional quantum Fourier transform can act as the locking operator in simultaneous teleportation of qu$N$its.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

Acknowledgments

The authors are grateful to the anonymous referee for invaluable comments and suggestions that help us improve the quality of the paper. This work is supported in part by the National Natural Science Foundation (Nos. 61272058, 61572532 and 61502179), the Natural Science Foundation of Guangdong Province of China (Nos. 10251027501000004 and 2014A030310265), the Research Foundation for the Doctoral Program of Higher School of Ministry of Education of China (No. 20100171110042), FCT project UID/EEA/50008/2013 and IT initiatives Probabilistic, Quantum and Differential Reasoning (PQDR) and Capacity and Privacy with Quantum Continuous Variables (CaPri).

Appendix A. Combination of the Two Strategies

In this appendix, we combine in one protocol the two strategies described in the above two sections: (i) if Bob and Charlie have the ability to intercept and resend the qubits going through the quantum channel, Alice can detect such behavior and interrupt the protocol; (ii) if Bob and Charlie are spatially separated, they can fairly decode their respective messages simultaneously through communication.

In the initialization phase, Alice, Bob and Charlie share two pair of entangled qu$N$its, denoted as

$$|\phi(00)\rangle_{A_1B} \otimes |\phi(00)\rangle_{A_2C}. \tag{33}$$

Alice and Bob also share classical information of the control qubits: $\{C\text{pos}_1(j)|j = 1, 2, \ldots, s\}$ and $\{C\text{state}_1(j)|j = 1, 2, \ldots, s\}$. Alice and Charlie also share classical information $\{C\text{pos}_2(j)|j = 1, 2, \ldots, s\}$ and $\{C\text{state}_2(j)|j = 1, 2, \ldots, s\}$.

The protocol of SDC of classical messages (b_1, b_2) and (c_1, c_2) works as follows:

(1) *Encoding.* Alice performs unitary operators $U(b_1 b_2)$ on qu$N$it A_1 and $U(c_1 c_2)$ on qu$N$it A_2 to encode (b_1, b_2) and (c_1, c_2), respectively.

(2) *Locking.* Alice joins the s control qubits with n message qubits $A_{1,1} \ldots A_{1,n}$, to form the ordered set of $(n + s)$ qubits $A_{1,1} \ldots A_{1,n+s}$. The jth control qubit is prepared in $C\text{state}_1(j) \in \mathcal{S}$ and is in the position $C\text{pos}_1(j)$, while the relative positions of the n message qubits are the same as before. Analogously, she forms the set $A_{2,1} \ldots A_{2,n+s}$.

For $j = 1$ to $n + s$, Alice applies the locking operator on qubits $A_{1,j} A_{2,j}$.

(3) *Communication.* Alice joins the r detect qubits with $(n + s)$ message and control qubits $A_{1,1} \ldots A_{1,n+s}$, to form the ordered set of $(n + s + r)$ qubits $A_{1,1} \ldots A_{1,n+s+r}$. The jth detect qubit is prepared in $D\text{state}_1(j) \in \mathcal{S}$ and is in the position $D\text{pos}_1(j)$, while the relative positions of the remaining $(n + s)$ qubits are the same as before. Analogously, she forms the set $A_{2,1} \ldots A_{2,n+s+r}$.

Alice sends $(n + s + r)$ qubits $A_{1,1} \ldots A_{1,n+s+r}$ to Bob, and $(n + s + r)$ qubits $A_{2,1} \ldots A_{2,n+s+r}$ to Charlie.

Alice waits for Bob and Charlie's acknowledgments of receiving all the $2(n + s + r)$ qubits. After they have sent their acknowledgments through unjammable classical communication channel, Alice sends $\{D\text{pos}_1(j)|j = 1, 2, \ldots, r\}$ and the bases of $\{D\text{state}_1(j)|j = 1, 2, \ldots, r\}$ to Bob. Analogously, Alice sends $\{D\text{pos}_2(j)|j = 1, 2, \ldots, r\}$ and the bases of $\{D\text{state}_2(j)|j = 1, 2, \ldots, r\}$ to Charlie.

For each $j = D\text{pos}_1(k)$, Bob measures qubit $A_{1,j}$ in either X or Z basis, according to $D\text{state}_1(k)$, and returns the measurement result to Alice. If the measurement result is not equal to $D\text{state}_1(k)$, Alice announces that a cheating behavior has been detected and stops the protocol.

Analogously, for each $j = D\text{pos}_2(k)$, Charlie measures qubit $A_{2,j}$ in either X or Z basis, according to $D\text{state}_2(k)$, and returns the measurement result to Alice. If the measurement result is not equal to $D\text{state}_2(k)$, Alice announces that a cheating behavior has been detected and stops the protocol.

Now the remaining $(n + s)$ received qubits at Bob's site form the ordered set $A_{1,1} \ldots A_{1,n+s}$, and the remaining $(n + s)$ received qubits at Charlie's site form the ordered set $A_{2,1} \ldots A_{2,n+s}$.

(4) *Unlocking.* For $j = 1$ to $n + s$, Bob sends qubit $A_{1,j}$ to Charlie, and then Charlie returns $A_{1,j}$ to Bob after performing the unlocking operator on qubits $A_{1,j} A_{2,j}$ at his site.

If $\exists k, C\text{pos}_1(k) = j$, Bob measures qubit $A_{1,j}$ in either X or Z basis, according to $C\text{state}_1(k)$. If the measurement result does not match $C\text{state}_1(k)$, he knows that Charlie did not return the real $A_{1,j}$ and stops the unlocking stage.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

Analogously, if $\exists k, C\text{pos}_2(k) = j$, Charlie measures qubit $A_{2,j}$ in either X or Z basis, according to $C\text{state}_2(k)$. If the measurement result does not match $C\text{state}_2(k)$, he knows that Bob did not send the real $A_{1,j}$ and stops the unlocking stage.

Now the remaining n received qubits at Bob's site form the ordered set $A_{1,1} \ldots A_{1,n}$, and the remaining n received qubits at Charlie's site form the ordered set $A_{2,1} \ldots A_{2,n}$.

(5) *Decoding.* Bob and Charlie measure qu$N$its A_1B and qu$N$its A_2C in the Bell basis respectively to achieve (b_1, b_2) and (c_1, c_2).

References

1. A. Holevo, *Probl. Inf. Transm.* **9**(3) (1973) 177.
2. C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69** (1992) 2881.
3. R. Horodecki, P. Horodecki, M. Horodecki and K. Horodecki, *Rev. Mod. Phys.* **81** (2009) 865.
4. X. S. Liu, G. L. Long, D. M. Tong and Feng Li, *Phys. Rev. A* **65** (2002) 022304.
5. A. Grudka and A. Wójcik, *Phys. Rev. A* **66** (2002) 014301.
6. Q.-B. Fan and S. Zhang, *Phys. Lett. A* **348**(3–6) (2006) 160.
7. A. Barenco and A. K. Ekert, *J. Mod. Opt.* **42**(6) (1995) 1253.
8. P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland and W. K. Wootters, *Phys. Rev. A* **54** (1996) 1869.
9. M. Ziman and V. Bužek, *Phys. Rev. A* **62** (2000) 052301.
10. J.-C. Hao, C.-F. Li and G.-C. Guo, *Phys. Lett. A* **278**(3) (2000) 113.
11. G. Bowen, *Phys. Rev. A* **63** (2001) 022302.
12. A. K. Pati, P. Parashar and P. Agrawal, *Phys. Rev. A* **72** (2005) 012329.
13. S. Mozes, J. Oppenheim and B. Reznik, *Phys. Rev. A* **71** (2005) 012311.
14. S. Wu, S. M. Cohen, Y. Sun and R. B. Griffiths, *Phys. Rev. A* **73** (2006) 042311.
15. P. S. Bourdon, E. Gerjuoy, J. P. McDonald and H. T. Williams, *Phys. Rev. A* **77** (2008) 022305.
16. M. R. Beran and S. M. Cohen, *Phys. Rev. A* **79** (2009) 032307.
17. E. Gerjuoy, H. T. Williams and P. S. Bourdon, *Phys. Rev. A* **79** (2009) 042315.
18. C.-Y. Huang, I.-C. Yu, F.-L. Lin and L.-Y. Hsu, *Phys. Rev. A* **79** (2009) 052306.
19. H. Situ, *Int. J. Theor. Phys.* **52**(10) (2013) 3779.
20. S. Bose, V. Vedral and P. L. Knight, *Phys. Rev. A* **57** (1998) 822.
21. H.-J. Lee, D. Ahn and S. W. Hwang, *Phys. Rev. A* **66** (2002) 024304.
22. Y. Yeo and W. K. Chua, *Phys. Rev. Lett.* **96** (2006) 060502.
23. D. Bruß, M. Lewenstein, A. Sen(De), U. Sen, G. M. D'Ariano and C. Macchiavello, *Int. J. Quantum Inf.* **04**(03) (2006) 415.
24. P. Agrawal and A. Pati, *Phys. Rev. A* **74** (2006) 062320.
25. X.-W. Wang, Y.-G. Shan, L.-X. Xia and M.-W. Lu, *Phys. Lett. A* **364**(1) (2007) 7.
26. L. Li and D. Qiu, *J. Phys. A, Math. Theor.* **40**(35) (2007) 10871.
27. S. Muralidharan and P. K. Panigrahi, *Phys. Rev. A* **77** (2008) 032321.
28. J.-C. Hao, C.-F. Li and G.-C. Guo, *Phys. Rev. A* **63** (2001) 054301.
29. C. L. Luo and X. F. Ouyang, *Int. J. Quantum Inf.* **07**(01) (2009) 365.
30. G. Q. Huang and C. L. Luo, *Int. J. Quantum Inf.* **07**(06) (2009) 1241.
31. H. Situ, *Int. J. Theor. Phys.* **53**(3) (2014) 1003.
32. M. Y. Wang, F. L. Yan, T. Gao and Y. C. Li, *Int. J. Quantum Inf.* **06**(01) (2008) 201.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

33. H. Situ and D. Qiu, *J. Phys. A, Math. Theor.* **43**(5) (2010) 055301.
34. S. Even and Y. Yacobi, Relations among public key signature systems, Technical Report, Technicon (1980).
35. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters, *Phys. Rev. Lett.* **70** (1993) 1895.
36. M. J. Fischer, N. A. Lynch and M. S. Paterson, *J. ACM* **32**(2) (1985) 374.
37. N. Paunković, J. Bouda and P. Mateus, *Phys. Rev. A* **84** (2011) 062331.
38. N. Asokan, M. Schunter and M. Waidner, Optimistic protocols for fair exchange, in *Proc. 4th ACM Conf. Computer and Communications Security (CCS '97)* (ACM, New York, NY, USA, 1997), pp. 7–17.
39. M. Ben-Or, O. Goldreich, S. Micali and R. L. Rivest, A fair protocol for signing contracts (extended abstract), in *Proceedings of the 12th Colloquium on Automata, Languages and Programming*, ed. W. Brauer, *ICALP*, Vol. 194, Lecture Notes in Computer Science (Springer, New York, 1985), pp. 43–52.
40. C. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proc. IEEE Int. Conf. Computers, Systems and Signal Processing* (IEEE Press, New York, 1984), pp. 175–179.
41. A. J. Menezes, S. A. Vanstone and P. C. V. Oorschot, *Handbook of Applied Cryptography*, 1st edn. (CRC Press, Boca Raton, FL, USA, 1996).

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by FLINDERS UNIVERSITY on 02/03/16. For personal use only.

International Journal of Quantum Information
Vol. 13, No. 6 (2015) 1550050 (10 pages)
© World Scientific Publishing Company
DOI: 10.1142/S0219749915500501

World Scientific
www.worldscientific.com

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

Quantum walk public-key cryptographic system

C. Vlachou*, J. Rodrigues†, P. Mateus‡, N. Paunković§
and A. Souto¶

Departamento de Matemática,
Instituto Superior Técnico, Universidade de Lisboa,
SQIG – Security and Quantum Information Group,
Instituto de Telecomunicações,
Avenida Rovisco Pais 1049-001, Lisboa, Portugal
*chrysoula.vlachou@ist.utl.pt
†jrodrigues@math.ist.utl.pt
‡pmat@math.ist.utl.pt
§n.paukovic@math.ist.utl.pt
¶a.souto@math.ist.utl.pt

Received 13 July 2015
Accepted 18 October 2015
Published 17 December 2015

Quantum Cryptography is a rapidly developing field of research that benefits from the properties of Quantum Mechanics in performing cryptographic tasks. Quantum walks are a powerful model for quantum computation and very promising for quantum information processing. In this paper, we present a quantum public-key cryptographic system based on quantum walks. In particular, in the proposed protocol the public-key is given by a quantum state generated by performing a quantum walk. We show that the protocol is secure and analyze the complexity of public key generation and encryption/decryption procedures.

Keywords: Quantum walks; public-key cryptographic systems; Holevo bound.

1. Introduction

Since the invention of writing, the need for secret/secure communication resulted in the development of cryptography — the art of "hidden communication". It started by using simple symbols as code words and evolved to the stage where the security is based on various mathematical hardness assumptions: A widely used RSA-based cryptographic system[1] relies on the conjecture that factoring large numbers is not feasible using standard computers, while the alternative lattice-based public-key cryptographic system[2] is based on the assumed difficulty of the so-called "shortest and closest vector problems" (also related to the well known $P \neq NP$ conjecture[3]).

¶Corresponding author.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

With the advent of quantum computation, and in particular after the discovery of the celebrated Shor's algorithm for factoring,[4] the security of most cryptographic systems currently in use became jeopardized and as a consequence the need for new cryptographic systems resilient to quantum adversaries arose. The above mentioned lattice-based cryptographic system is resilient to particular quantum adversaries that execute Shor's algorithm, as its security is not based on the factoring problem; nevertheless, it does rely on another mathematical assumption, thus leaving open the question of whether it is more than just computationally secure. A completely different approach is the one that takes advantage of quantum mechanics and assumes that both sender and receiver have access to a quantum computer. The major advantage of this approach is that the security, rather than relying on mathematical/computational hardness assumptions, is now based on the laws of quantum mechanics.

Quantum cryptography has been among hot subjects of research in the last decades. It was firstly considered by Wiesner in the late 60s and early 70s, who introduced the notions of quantum multiplexing and money (which was only published a decade later, see Ref. 5), and further developed by Bennet and Brassard[6] in their famous Quantum Key Distribution (QKD) BB84 protocol. Subsequently, Mayers showed the unconditional security of BB84 protocol i.e. the security that is based on the laws of (quantum) physics (see Ref. 7). Rapid development of experimental techniques resulted in the implementation of QKD, see for example, Ref. 8, and nowadays one can even buy devices implementing it on the online market from Clavis QKD Platform.[9]

As an alternative to QKD, another way for two parties to communicate privately has been developed, using quantum public-key cryptosystems: in Ref. 10 the authors propose a scheme based on quantum trapdoor one-way functions, and in Ref. 11 a system is presented whose security is based on (quantum) computational indistinguishability of quantum states. More recently, Nikolopoulos[12] presented a secure public-key cryptographic system based on single-qubit rotations.

In this paper, we propose a new quantum public-key cryptographic system in which the public keys are quantum states generated by running a quantum walk, rather than by performing single-qubit rotations, while the private key consists of: (i) the quantum walk, (ii) the number of steps of the walk and (iii) the initial state. Regarding the use of quantum walks in cryptography, see for example the recent work by Yan *et al.*[13]

The paper is organized as follows: in the next section, we start with basic definitions and notation used throughout the paper. In Sec. 3, we present our public-key cryptographic system and discuss its security and efficiency. In Sec. 4, we summarize the results obtained and point some future research direction.

2. Quantum Walks

The concept of random walk was introduced to describe the behavior of a "walker" over a path who, at each step, can choose to follow one of the possible directions with

a certain *a priori* fixed probability. It was shown to be very useful in computer science (sampling massive online graphs, image segmentation, estimating the size of the World Wide Web, wireless networking, etc.), physics (modeling the Brownian motion, studying polymers, etc.), and many other fields of research (financial economics, medicine and biology, psychology, etc.).

Quantum walks are the quantum counterpart of random walks and were first considered in Ref. 14. Unlike the classical case, in which the state of the walker is described by a probability distribution over the allowed positions, in the quantum scenario the state of the walker is given by a superposition of positions. Quantum walks are a successful tool in algorithmic theory since they provide an exponential speedup over classical computations for several oracular problems[15–17] and polynomial speedup for many important classical solutions of problems such as determining whether all the elements of a list are distinct,[18] and in search problems in general.[19] Moreover, they are an important computational primitive, since they permit universal quantum computation.[20,21] For practical implementations, see for example the recent book.[22] One can study different types of quantum walks, determined by their time evolution (discrete- versus continuous-time), underlying position space and their topology (walks on the line, lattice, circle, graphs, etc.). In this work we use discrete-time quantum walks on the circle.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

2.1. *One-dimensional discrete-time quantum walks*

2.1.1. *Basic dynamics*

In a discrete-time quantum walk on an infinite line, we consider the movement of a walker along discrete positions on it, labeled $i \in \mathbb{Z}$. At each step the particle moves to the left or to the right, depending on the state of the internal degree of freedom, commonly called the coin state. The position and the coin state of the particle are from the Hilbert spaces $\mathcal{H}_p = span\{|i\rangle : i \in \mathbb{Z}\}$ and $\mathcal{H}_c = span\{|R\rangle, |L\rangle\}$, respectively.

The evolution of the system at each step of the walk is described by the unitary operator:

$$\hat{U} = \hat{S}(\hat{I}_p \otimes \hat{U}_c). \tag{1}$$

In the above expression, $\hat{I}_p$ is the identity operator on $\mathcal{H}_p$, $\hat{S}$ is the shift operator

$$\hat{S} = \sum_{i\in\mathbb{Z}} (|i+1\rangle\langle i| \otimes |R\rangle\langle R| + |i-1\rangle\langle i| \otimes |L\rangle\langle L|) \tag{2}$$

and $\hat{U}_c \in U(2)$ is the coin operator acting on $\mathcal{H}_c$. The general expression for $\hat{U}_c$ is:

$$\hat{U}_c = \hat{U}_c(\theta, \xi, \zeta) = \begin{bmatrix} e^{i\xi}\cos\theta & e^{i\zeta}\sin\theta \\ -e^{-i\zeta}\sin\theta & e^{-i\xi}\cos\theta \end{bmatrix}. \tag{3}$$

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

2.1.2. *Shift operator on the circle*

To simulate the walk on a circle, one could either identify positions $-N$ and N of the above discussed line, or connect the two, thus altering the corresponding shift operator. In the former, the circle has an even number of positions ($2N$), while in the latter it has an odd number of positions ($2N+1$).

Without loss of generality, we consider the position Hilbert space $\mathcal{H}_p = span\{|i\rangle : i \in \{0, \ldots, N-1\}\}$. The general expression for a shift operator on a circle with N positions is:

$$\begin{aligned}\hat{S} &= \sum_{i=0}^{N-1}(|i+1(\mathrm{mod}\, N)\rangle\langle i| \otimes |R\rangle\langle R| + |i-1(\mathrm{mod}\, N)\rangle\langle i| \otimes |L\rangle\langle L|) \\ &= \hat{T}_1 \otimes |R\rangle\langle R| + \hat{T}_{-1} \otimes |L\rangle\langle L|, \end{aligned} \tag{4}$$

where

$$\hat{T}_m = \sum_{i=0}^{N-1} |i + m(\mathrm{mod}\, N)\rangle\langle i|, \tag{5}$$

is the m — position translation operator.

3. Public-Key Encryption Based on Discrete-Time Quantum Walks

We now present the public-key cryptographic system using quantum walks on a circle. To generate the public key, we use a discrete number of possible walks $\hat{U}_k = \hat{S}[\hat{I} \otimes \hat{U}_c(\theta_k, \xi_k, \zeta_k)]$, with $\theta_k = \xi_k = \zeta_k = k\frac{2\pi}{d}$, $k \in \mathcal{I} = \{1, 2, \ldots, d\}$ and $d \in \mathbb{N}$.

Protocol 1 (Public-key encryption scheme).
Inputs for the protocol

- Message to transfer: $m \in \{0, \ldots, 2^n - 1\}$, i.e. a message of at most n bits;
- Secret key $SK = (\hat{U}_k, t, l)$ where:
  $\hat{U}_k$ with $k \in \mathcal{I} = \{1, 2, \ldots, d\}$, $t \in \mathcal{T} = \{t_0, \ldots, t_{\max}\} \subset \mathbb{N}$ and $l \in \{0, \ldots, 2^n - 1\}$.

Public-key generation

- Alice chooses uniformly at random $l \in \{0, \ldots, 2^n - 1\}$ and $s \in \{L, R\}$, and generates the initial state $|l\rangle|s\rangle$;
- Then she chooses, also uniformly at random, the walk $\hat{U}_k = \hat{S}(\hat{I}_p \otimes \hat{U}_c)$ and the number of steps $t \in \mathcal{T}$;
- Finally, she generates the public key:

$$|\psi_{PK}\rangle = \hat{U}_k^t|l\rangle|s\rangle = [\hat{S}(\hat{I}_p \otimes \hat{U}_c)]^t|l\rangle|s\rangle. \tag{6}$$

Message Encryption

- Bob obtains Alice's public key $|\psi_{PK}\rangle$;

- He encrypts m by applying spatial translation to obtain:

$$|\psi(m)\rangle = (\hat{T}_m \otimes \hat{I}_c)|\psi_{PK}\rangle; \tag{7}$$

- Bob sends $|\psi(m)\rangle$ to Alice.

Message Decryption

- Alice applies $\hat{U}_k^{-t}$ to the state $|\psi(m)\rangle$;
- She performs the measurement

$$\hat{M} = \sum_i |i\rangle\langle i| \otimes \hat{I}_c \tag{8}$$

and obtains the result m'. The message sent by Bob is $m = m' - l \pmod N$.

3.1. *Correctness of the protocol*

Proposition 1. *The above protocol is correct, in the sense that if Alice and Bob follow it, and no third party intervenes during its execution, at the end of the decryption phase Alice recovers the message sent by Bob with probability* 1.

Proof. The correctness of the protocol when Alice and Bob follow the prescribed steps is a direct consequence of the fact that the quantum walk $\hat{U}_k^t$ commutes with any translation $\hat{T}_m$. Thus, the state of the system before the final step of the decryption phase (measurement), is:

$$\begin{aligned} |\psi_f\rangle &= \hat{U}_k^{-t}|\psi(m)\rangle \\ &= \hat{U}_k^{-t}(\hat{T}_m \otimes \hat{I}_c)\hat{U}_k^t|l\rangle|s\rangle \\ &= (\hat{T}_m \otimes \hat{I}_c)|l\rangle|s\rangle \\ &= |l + m(\mathrm{mod}\, N)\rangle|s\rangle. \end{aligned} \tag{9}$$

Hence, upon measuring $\hat{M}$ and obtaining $m' = l + m(\mathrm{mod}\, N)$, the last modular operation performed in the last step of the *Message Decryption* reveals that the decrypted message is indeed m. □

In the previous Proposition, the crucial point is the fact that the quantum walk commutes with the translation. Below, we prove in detail that $\hat{U}_k^t$ and $(\hat{T}_m \otimes \hat{I}_c)$ commute.

Lemma 1. *Let $N \geq 2^n$ where n is a fixed integer. Let $\hat{U}_k^t$ be a quantum walk from Protocol 1 and let $\hat{T}_m$ denote the translation operator for m positions modulo N. Then $\hat{U}_k^t$ and $(\hat{T}_m \otimes \hat{I}_c)$ commute.*

Proof. Notice that the action of any $\hat{U}_k$ used in Protocol 1 can be written as:

$$\hat{U}_k|l\rangle|s\rangle = \alpha_{L(s)}|l-1\rangle|L\rangle + \alpha_{R(s)}|l+1\rangle|R\rangle, \tag{10}$$

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

where $|L\rangle$ and $|R\rangle$ are the orthogonal coin states and $\alpha_{L/R(s)}$ is the probability amplitude to find the walker in position $l-1$ or $l+1$, depending on its spin. Notice also that $\hat{T}_m$ is defined as:

$$\hat{T}_m|l\rangle = |l + m(\mathrm{mod}\, N)\rangle. \tag{11}$$

Then, for any element of the form $|l\rangle|s\rangle$ we have:

$$\begin{aligned}(\hat{T}_m \otimes \hat{I}_c)\hat{U}_k|l\rangle|s\rangle &= (\hat{T}_m \otimes \hat{I}_c)[\alpha_{L(s)}|l-1\rangle|L\rangle + \alpha_{R(s)}|l+1\rangle|R\rangle] \\ &= \alpha_{L(s)}|l-1+m(\mathrm{mod}\, N)\rangle|L\rangle \\ &\quad + \alpha_{R(s)}|l+1+m(\mathrm{mod}\, N)\rangle|R\rangle. \end{aligned} \tag{12}$$

On the other hand, we also have:

$$\begin{aligned}\hat{U}_k(\hat{T}_m \otimes \hat{I}_c)|l\rangle|s\rangle &= \hat{U}_k|l+m(\mathrm{mod}\, N)\rangle|s\rangle \\ &= \alpha_{L(s)}|l-1+m(\mathrm{mod}\, N)\rangle|L\rangle \\ &\quad + \alpha_{R(s)}|l+1+m(\mathrm{mod}\, N)\rangle|R\rangle. \end{aligned} \tag{13}$$

□

Observe that this lemma can be extended to more general shift operations, which allow jumps for two or more positions, or leave position state unchanged, depending on the coin state.

3.2. *Security of the protocol*

The protocol consists of two phases. In the first, Alice sends a public key $|\psi_{PK}\rangle$ to Bob. In the second, upon encrypting the message m, Bob sends back the state $|\psi(m)\rangle$ to Alice. Therefore, one has to show the security of the secret key during the first phase and the security of the message during the second phase.

Our proof of the security is based on Holevo's Theorem, that bounds the amount of classical information that an eavesdropper can retrieve from a given quantum mixed state by means of a POVM measurement.

Let us denote by $\hat{\rho}_{PK}$ the mixed state of the public key, as perceived by Eve, who does not know *a priori* the secret key SK chosen by Alice. Even if Eve were to know $\hat{U}_k$ and t, $\hat{\rho}_{PK}$ is completely mixed:

$$\begin{aligned}\hat{\rho}_{PK} &= \hat{U}_k^t \left[\frac{1}{2^{n+1}} \sum_{l=0}^{2^n-1} \sum_{s\in\{L,R\}} |l\rangle\langle l| \otimes |s\rangle\langle s|\right] (\hat{U}_k^t)^\dagger \\ &= \hat{U}_k^t \left(\frac{1}{2^{n+1}} \hat{I}_p \otimes \hat{I}_c\right) (\hat{U}_k^t)^\dagger \\ &= \frac{1}{2^{n+1}} \hat{I}_p \otimes \hat{I}_c. \end{aligned} \tag{14}$$

Assuming that Eve performs a measurement on $\hat{\rho}_{PK}$, Holevo's Theorem implies that the mutual information $I(SK, E)$ between the secret key SK and her inference E is

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

bounded from above by the von Neumann entropy of this state:

$$I(SK, E) \leq S(\hat{\rho}_{PK}) = -\mathrm{Tr}(\hat{\rho}_{PK} \log \hat{\rho}_{PK}) = n + 1. \tag{15}$$

To conclude that the protocol is secure we have to show that the mutual information is very small compared to the Shannon entropy of the secret key. Indeed, the Shannon entropy of the secret key depends on the probability to choose $\hat{U}_k$, t and the initial state $|l\rangle|s\rangle$. In the following we denote by p_k the probability to choose $\hat{U}_k$ from the set $\{\hat{U}_k | k \in \mathcal{I} = \{1, 2, \ldots, d\}\}$, by p_t the probability to run the walk for t steps, with $t \in \mathcal{T} = \{t_0, \ldots, t_{\max}\}$ and by $p_{l,s}$ the probability to choose $|l\rangle|s\rangle$ as the initial state, where $l \in \{0, 1, \ldots, 2^n - 1\}$ and $s \in \{L, R\}$. Since this choices are random and independent, the probability of a certain secret key SK is given by:

$$p_{SK} = p_k \, p_t \, p_{l,s} = \frac{1}{d \, |\mathcal{T}| \, 2^{n+1}}, \tag{16}$$

where $|\mathcal{T}|$ is the cardinality of $\mathcal{T}$.

The above probability distributions are uniform, so the Shannon entropy of the secret key is:

$$\begin{aligned} H(p_{SK}) &= -\sum_{k \in \mathcal{I}} \sum_{t \in \mathcal{T}} \sum_{l=0}^{2^n - 1} \sum_{s \in \{L,R\}} p_k \, p_t \, p_{l,s} \log_2(p_k \, p_t \, p_{l,s}) \\ &= \log_2(d \, |\mathcal{T}| \, 2^{n+1}) \\ &= \log_2(d \, |\mathcal{T}|) + n + 1. \end{aligned} \tag{17}$$

Thus, we have:

$$I(SK, E) \leq S(\hat{\rho}_{PK}) < H(p_{SK}). \tag{18}$$

With the appropriate choice of $|\mathcal{T}|$ and d, e.g. $|\mathcal{T}|, \log d \approx poly(n)$, for sufficiently large n, the Shannon entropy of the secret key has a polynomial overhead over the von Neumann entropy of the public key as seen by Eve,

$$H(p_{SK}) - S(\hat{\rho}_{PK}) = \log_2(d \, |\mathcal{T}|) \approx poly(n). \tag{19}$$

This way, upon obtaining the maximal possible information about the secret key, given by $S(\hat{\rho}_{PK})$, Eve's uncertainty (in the number of bits) of the SK is still polynomial in n, i.e. the number of keys consistent with the information she has is exponential in n. We note that the choice of $d \approx \exp(n)$ secures the secrecy of the encrypted message, while $|\mathcal{T}| \approx poly(n)$ was chosen to maintain the protocol's efficiency, discussed in the next section.

In the rest of this section, we discuss the security of the message m during the second phase of the protocol, when Bob sends the encrypted message $|\psi(m)\rangle = (\hat{T}_m \otimes \hat{I}_c)|\psi_{PK}\rangle$ to Alice. Without knowing the secret key, the state perceived by Eve is still a complete mixture:

$$\hat{\rho}_E = (\hat{T}_m \otimes \hat{I}_c)\left(\frac{1}{2^{n+1}} \hat{I}_p \otimes \hat{I}_c\right)(\hat{T}_m \otimes \hat{I}_c)^\dagger = \frac{1}{2^{n+1}} \hat{I}_p \otimes \hat{I}_c. \tag{20}$$

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

The most that Eve can learn is the very quantum state $|\psi(m)\rangle$ (although, as proven above, even that is impossible, unless with negligible probability). Nevertheless, without knowing the secret key, this information is not enough for Eve to infer the message encrypted by Bob. This is a simple consequence of the fact that for each allowed encryption state, there exists a suitably chosen secret key that can decrypt *any* message m. Indeed, a state $|\psi(m)\rangle$ that for the secret key $SK = (\hat{U}_k, t, l)$ corresponds to the message m, for the secret key $SK' = (\hat{U}_k, t, l - \Delta l)$ corresponds to the message $m + \Delta l$ (below, the subscripts SK and SK' explicitly denote the secret key used to encrypt the corresponding messages m and $m + \Delta l$, respectively):

$$\begin{aligned} |\psi(m)\rangle_{SK} &= (\hat{T}_m \otimes \hat{I}_c)|\psi_{PK}\rangle \\ &= (\hat{T}_m \otimes \hat{I}_c)\,\hat{U}_k^t(\hat{T}_{\Delta l} \otimes \hat{I}_c)|l - \Delta l\rangle|s\rangle \\ &= (\hat{T}_{m+\Delta l} \otimes \hat{I}_c)\,\hat{U}_k^t|l - \Delta l\rangle|s\rangle \\ &= |\psi(m + \Delta l)\rangle_{SK'}. \end{aligned} \tag{21}$$

3.3. *Efficiency of the protocol*

In this section, we show the efficiency of the proposed protocol, i.e. the overall time τ required for its execution (public-key generation, message encryption and message decryption) scales polynomially with the length n of the message.

The public-key generation as well as the message decryption are efficient procedures, since the quantum walks used to perform them are efficient. Indeed, denoting by $\Delta\tau_w$ the time required for a single step $\hat{U}$ of the walk, the full walk $\hat{U}^t$ is completed in time $\tau = t \cdot \Delta\tau_w$. In the previous section we took $t \approx poly(n)$ for security purposes, a choice which is also adequate for the efficiency of the quantum walk: the time required to perform the walk is polynomial in n.

In addition to this, for the overall protocol to be efficient, the message encryption, given by the translation operator $\hat{T}_m$, has to be efficient as well. It might seem at first that the encryption of the message is not efficient, as it requires $\mathbb{O}(2^n)$ single-position translations, $\hat{T}_m = (\hat{T}_1)^m$. Below, we show that this is not necessarily a non-efficient procedure, i.e. various practical implementations of $\hat{T}_m$ are indeed efficient.

In case the system that performs the quantum walk consists of $n + 1$ qubits (n carrying the position of the walker plus the coin one), such that the states of the computational basis encode different positions (see for example, Ref. 23), the translation operator $\hat{T}_m$ is nothing but the addition by m, which is an efficient operation for a quantum computer. Alternatively, in those cases of physical realizations in which different position states $|i\rangle$ are given by distinct spatial positions (see for example, implementations based on integrated photonics[24]), Bob can simply relabel the positions on the device that carries the quantum state of the public key, i.e. $i \to i - m$, which is also efficient, as he can do it in parallel at the same time for all position states.

4. Conclusions

In this paper, we presented a quantum public-key cryptographic system based on quantum walks. Unlike the recently proposed protocol[12] that uses single-qubit rotations to generate the public key, in our scheme the execution of a quantum walk, in general, results in entangled quantum states as public keys, thus increasing the practical security of our scheme (an eavesdropper has to, in general, perform more complex operations to extract information from entangled rather than from product states). Using Holevo's theorem, we proved the protocol's security. We also analyzed the complexity of our public-key generation and message encryption/decryption and showed their efficiency, i.e. that the complexity of our protocol scales polynomially with the size of the message.

Future research includes designing other security protocols based on the use of quantum walks, such as oblivious transfer (along the lines of the recently proposed protocol presented in Ref. 25), commitment schemes and other privacy functionalities.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

Acknowledgments

This work was partially supported, under the CV-Quantum initiative at IT, by FCT PEst-OE/EEI/LA0008/2013 and UID/EEA/50008/2013, namely via the FCT CaPri initiative. We also acknowledge bilateral scientific co-operation between Portugal and Serbia through the project "Noise and measurement errors in multi-party quantum security protocols", No. 451-03-01765/2014-09/04. C. Vlachou and J. Rodrigues and A. Souto acknowledge, respectively, the FCT grants PD/BD/52652/2014, SFRH/BD/75085/2010 and SFRH/BPD/76231/2011. N. Paunković acknowledges discussions with D. Arsenović, D. Popović and S. Prvanović.

References

1. R. L. Rivest, A. Shamir and L. Adleman, *Commun. ACM* **21**(2) (1978) 120.
2. O. Goldreich, S. Goldwasser and S. Halevi, Public-key cryptosystems from lattice reduction problems, in *Proc. 17th Annual Int. Cryptology Conf. Advances in Cryptology, CRYPTO '97* (Springer-Verlag, London, UK, 1997), pp. 112–131.
3. L. Fortnow, *The Golden Ticket — P, NP, and the Search for the Impossible* (Princeton University Press, New Jersey, 2013).
4. P. Shor, *SIAM J. Comput.* **26**(5) (1997) 1484.
5. S. Wiesner, *SIGACT News* **15**(1) (1983) 78.
6. C. Bennett and G. Brassard, Quantum Cryptography: Public Key Distribution and Coin Tossing, in *Proc. IEEE Int. Conf. Computers, Systems and Signal Processing* (IEEE Press, New York, 1984), pp. 175–179.
7. D. Mayers, *J. ACM* **48**(3) (2001) 351.
8. T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. Rarity, A. N. Zeilinger and H. Weinfurter, *Phys. Rev. Lett.* **98** (2007) 010504.

9. IdQuantic. Clavis QKD Platform, http://www.idquantique.com/photon-counting/clavis-qkd-platform/, 2015.
10. T. Okamoto, K. Tanaka and S. Uchiyama, Quantum public-key cryptosystems, in *Proc. of CRYPT0 2000* (Springer-Verlag, Berlin-Heidelberg, 2000), pp. 147–165.
11. A. Kawachi, T. Koshiba, H. Nishimura and T. Yamakami, in *Proc. Advances in Cryptology — EUROCRYPT 2005, 24th Annual Int. Conf. on the Theory and Applications of Cryptographic Techniques*, Lecture Notes in Computer Science, Vol. 3494 (Springer, Berlin-Heidelberg, 2005), pp. 268–284.
12. G. Nikolopoulos, *Phys. Rev. A* **77** (2008) 032348.
13. Y. Yang, Q. Pan, S. Sun and P. Xu, *Sci. Rep.* **5** (2015) 7784.
14. Y. Aharonov, L. Davidovich and N. Zagury, *Phys. Rev. A* **48** (1993) 1687.
15. A. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann and D. Spielman, Exponential algorithmic speedup by a quantum walk, in *Proc. 35th Annual ACM Symposium on Theory of Computing*, STOC '03 (ACM, New York, 2003), pp. 59–68.
16. J. Kempe, *Contemp. Phys.* **44**(4) (2003) 307.
17. A. Ambainis, *Int. J. Quantum Inform.* **1**(4) (2003) 507.
18. M. Santha, Quantum walk based search algorithms, in *Theory and Applications of Models of Computation*, eds. M. Agrawal, D. Du, Z. Duan and A. Li, Lecture Notes in Computer Science, Vol. 4978 (Springer, Berlin-Heidelberg, 2008), pp. 31–46.
19. R. Portugal, *Quantum Walks and Search Algorithms*, Quantum Science and Technology (Springer, New York, 2013).
20. N. B. Lovett, S. Cooper, M. Everitt, M. Trevers and V. Kendon, *Phys. Rev. A* **81** (2010) 042330.
21. A. M. Childs, D. Gosset and Z. Webb, *Science* **339**(6121) (2013) 791.
22. K. Manouchehri and J. Wang, *Physical Implementation of Quantum Walks* (Springer, Berlin, 2013).
23. C. Ryan, M. Laforest, J. Boileau and R. Laflamme, *Phys. Rev. A* **72** (2005) 062317.
24. L. Sansoni, F. Sciarrino, G. Vallone, P. Mataloni, A. Crespi, R. Ramponi and R. Osellame, *Phys. Rev. Lett.* **108** (2012) 010502.
25. P. Mateus, N. Paunkovic, J. Rodrigues and A. Souto, Enhancing privacy with quantum networks, in *Proc. CMS 2014*, Lecture Notes in Computer Science, Vol. 8735 (Springer, Berlin-Heidelberg, 2014), pp. 147–153.

Int. J. Quantum Inform. Downloaded from www.worldscientific.com
by WEIZMANN INSTITUTE OF SCIENCE on 01/02/16. For personal use only.

International Journal of Quantum Information
Vol. 6, No. 2 (2008) 281–302
© World Scientific Publishing Company

World Scientific
www.worldscientific.com

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

ON THE POWER OF QUANTUM TAMPER-PROOF DEVICES*

J. BOUDA

Faculty of Informatics, Masaryk University,
Botanická 68a, 602 00 Brno, Czech Republic
bouda@fi.muni.cz

P. MATEUS†, N. PAUNKOVIC‡ and J. RASGA§

SQIG – Instituto de Telecomunicações and IST–TU Lisbon,
Av. Rovisco Pais, 1049-001, Lisbon, Portugal
†*pmat@math.ist.utl.pt*
‡*npaunkov@math.ist.utl.pt*
§*jfr@math.ist.utl.pt*

Received 20 February 2008

We show how quantum tamper-proof devices (QTPD's) can be used to attack and to develop security protocols. On one hand, we prove that it is possible to transfer proofs of zero-knowledge protocols using QTPD's. This attack can be extended to other security schemes where privacy is important. On the other hand, we present a fair contract signing protocol using QTPD's where there is no communication with Judge during the exchange phase (which is impossible classically). In the latter case, we make use of decoherence in the quantum state of the QTPD to implement a global clock over the asynchronous network. QTPD's seem to be possible to implement with existing quantum hardware, due to the fact that it is hard to isolate quantum memory from interference. These theoretical results contribute to justify the implementation of QTPD's.

Keywords: Quantum tamper-proof device; zero-knowledge; contract signing.

1. Introduction

Recently, research on quantum hardware and software has been attracting a lot of attention, especially concerning applications in security and cryptography. On one hand, Shor showed that when a full-fledged quantum computer is available, all cryptosystems based on factorization and discrete logarithms will fail. On the other hand, quantum systems are more secure than classical ones due to the fact that it is not possible to *determine* a completely unknown quantum state or, in other words, to clone an unknown quantum state. Indeed, it is possible to establish a

*This paper was presented at the *LQCIL'07 Workshop on Quantum Cryptography and Security*, which take place in 18–20 July 2007, in the premises of Instituto Superior Técnico (IST), Technical University of Lisbon, in Portugal.

perfectly secure secret key over a (authenticated) quantum channel which cannot be mimicked with classical channels.

In this paper, we focus on the applications of quantum tamper-proof devices (QTPD's). Indeed, quantum technology seems a very good candidate for developing tamper-proof devices due to its fragility and therefore, ease of detecting tampering. Moreover, quantum systems have an inbuilt pure random generator that is very useful in security. Herein, we show how QTPD's can be used, on one hand, to attack existing secure systems, and on the other hand, to implement protocols that are not possible classically. We show in detail how to attack with a QTPD the impossibility of transferring the proof for some class of zero-knowledge proof systems. Moreover, we present a fair contract signing protocol (which is impossible classically) using QTPD's.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

Roughly speaking, a zero-knowledge proof system[1] is a two party protocol, between a verifier and a prover, in which the prover wants to show to the verifier that he knows some secret without conveying any information about this secret to the verifier (*zero-knowledge property*). Zero-knowledge proof systems are used in several protocols, like identification protocols,[2] e-voting protocols[3] and e-money protocols.[4] While the zero-knowledge property is the most important security aspect these systems must fulfill, in the folklore, another security property is assumed to hold in zero-knowledge systems. This property is known as the *impossibility of transferring the proof*, and it states that at the end of the protocol, the verifier cannot prove to a third party that he had interacted with the prover. So, the impossibility of transferring the proof is crucial in satisfying several anonymity properties for more complex protocols like undeniable signatures, identification protocols and electronic elections. In identification protocols, the impossibility of transferring the proof allows the anonymity of the prover in the sense that the verifier cannot show to a third party that a certain prover has been identified to the verifier.

To prove the impossibility of transferring the proof in a particular zero-knowledge proof system, it is usually assumed that the verifier has complete control over his memory space and that he can roll back to a previous state whenever he needs. In practice, this constitutes a very generous assumption. In a more general context, the verifier may have restrictions on accessing his memory; for instance, he might be able to read/write his memory only via a QTPD. In this case, we are able to set up an attack where the verifier transfers the proof with a QTPD to a third party.

Fortunately, QTPD's can also be used to create new protocols, and not only to attack existing ones. We will show how a QTPD with decoherence allows one to set up a fair contract signing protocol.

Contract signing[5] is an important security task with many applications, namely to the stock market and others. In a nutshell, a contract signing protocol is a two-party protocol between Alice and Bob, who share a contract and want to exchange each other's signature of the contract. However, Alice and Bob do not trust each other, and thus the protocol must be *fair*. In other words, either both Alice and Bob

get each other's signature or neither of them does. A simple *fair* contract signing protocol is one where Alice and Bob do not communicate at all. Since neither of them receives the signature, the protocol is fair. For this reason, it is also important to require the protocol to be *viable*. In other words, if both parties behave honestly, then both of them will get each other's signature.

It has been shown that there exists no fair and viable contract signing protocol,[5] unless signing parties communicate with a trusted Judge in the signing phase. The proof is rather simple, and it is related to the impossibility of establishing a distributed consensus in asynchronous networks.[6] Indeed, if such a protocol exists, it has a final step where one message is exchanged, say, from Alice to Bob. In this case, Alice already has all the information she requires to compute Bob's signature of the contract (and Bob does not). Therefore, if she does not send the last message, the protocol reaches an unfair state. One way to get around this difficulty is to relax the fairness condition probabilistically. Probabilistic fairness allows one agent to have ϵ more probability of binding the contract over the other agent. In this case, solutions where the number of exchanged messages between the agents is minimized have been found in Ref. 7. Another way around is to consider optimistic protocols that do not require communication with Judge unless something wrong comes up.[8] The basic idea for this type of solution is to ask for the assistance of Judge when the contract signing protocol has been initiated and some message is missing.

Herein we propose a completely new solution for fair contract signing. We assume a QTPD with decoherence that works as a tamper-proof global clock in the asynchronous network. In this way, Judge is able to know *a posteriori* whether the parties are committed or not to the contract, upon a given time, without any communication (with Judge) *through the signing phase*.

The paper is organized as follows. In Sec. 2, we formalize QTPD's as quantum automata. In Sec. 3, we introduce the basic notion of zero-knowledge proof systems and the transference attack with QTPD's. In Sec. 4, we present the fair contract signing. Finally, we draw some conclusions in Sec. 5.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

2. Quantum Tamper-Proof Devices

In this section, we formalize the concept of a *quantum tamper-proof device*, i.e. a machine for which an evil agent has no possibility of arbitrarily changing its internal state. We describe a QTPD as a quantum automaton. The notion of a quantum automaton that we consider is a generalization of the notion introduced in Ref. 9, where we extend the set of possible outputs and allow measurements as inputs.

Definition 1. A *quantum tamper-proof device* is a quantum automaton, i.e. a tuple $Q = (I, \Gamma, O, H, |\psi_0\rangle, U, A)$, where:

- I is a finite set of *transition symbols*;
- Γ is a finite set of *observation symbols*;
- $O \subseteq \mathbb{R}$ is a finite set of *output symbols*;

- H is finite-dimensional Hilbert space called the *state space*;
- $|\psi_0\rangle \in H$ is a unit vector in H called the *initial state*;
- $U = \{U_i\}_{i\in I}$ is a family of unitary transformations in H called the *family of transitions*;
- $A = \{A_\gamma\}_{\gamma\in\Gamma}$ is a family of observables over H called the *family of observables* such that the spectrum of A_γ is contained in O for all $\gamma \in \Gamma$.

The dynamics of a quantum automaton is the following. Two types of inputs are possible: transition inputs I and observation inputs Γ. Transition inputs are associated with unitary transformations $\{U_i\}_{i\in I}$ while observation inputs are associated with observables $\{A_\gamma\}_{\gamma\in\Gamma}$. The automaton is a reactive machine, which changes state when an input is read and outputs a value when an observation input is read. The set of outputs O is a (finite) subset of the real numbers containing the spectrum of all observables A_γ of the automaton. The state of the automaton is a unit vector over the state space H, which is a space where all unitary transformations and observables are defined. The automaton starts at the initial state $|\psi_0\rangle$. The state of the automaton evolves deterministically as transition inputs are read and probabilistically when the inputs are observations, respecting the postulates of quantum mechanics. That is, when a transition input i is read, the state $|\psi\rangle$ of the automaton evolves to $U_i|\psi\rangle$ without outputting any value. If an observation input γ is read, the state $|\psi\rangle$ evolves to $P_o|\psi\rangle/\|P_o|\psi\rangle\|$ with probability $\|P_o|\psi\rangle\|$, for any eigenvalue $o \in O$ of A_γ, where P_o is the projection onto the eigenspace of A_γ associated with o. In this case, the automaton outputs o. As we shall see, we are usually not interested in unitary transitions, but only in observations.

To illustrate the concept, we describe a quantum automaton that generates pure random bits.

Example 1 (pure random bit generator). Consider the quantum automaton $(I, \Gamma, O, H, |\psi_0\rangle, U, A)$, where:

- $I = \{t\}$;
- $\Gamma = \{o\}$;
- $O = \{-1, 1\}$;
- $H = \mathbb{C}^2$;
- $|\psi_0\rangle = |0\rangle$;
- U_t is the Hadamard transformation;
- A_o is the observable $-|0\rangle\langle 0| + |1\rangle\langle 1|$.

Recall that the Hadamard transformation is the unitary transformation that maps $|0\rangle$ to $1/\sqrt{2}(|0\rangle+|1\rangle)$ and $|1\rangle$ to $1/\sqrt{2}(|0\rangle-|1\rangle)$. Each time the sequence $t\,o$ of inputs is read, the automaton generates a random bit. The bit generation works as follows. Initially, the input t associated with the Hadamard transformation is read, setting the state to $1/\sqrt{2}(|0\rangle + |1\rangle)$. Then the observation input o is read, outputting -1 or 1 with probability 1/2 each, and the state evolves to $|0\rangle$ or $|1\rangle$. For obtaining more random bits, another sequence $t\,o$ should be read. Observe that for the next

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

random bits, when t is read, the state evolves to $1/\sqrt{2}(|0\rangle \pm |1\rangle)$. Nevertheless, when any of these states are observed by A_o the output is -1 or 1 with probability $1/2$.

We will also need to deal with QTPD's where the internal state is decaying over time, in order to present the contract signing protocol. This feature is essential in modeling global clocks on the asynchronous network. For simplicity we will consider that the state is evolving over time until some measurement is performed. At that point the systems stops evolving and behaves as a common quantum automaton (Definition 1).

Definition 2. A *quantum tamper-proof device with decoherence* is a quantum automaton, i.e. a tuple $Q = (I, \Gamma, O, H, \rho, U, A)$ where:

- I is a finite set of *transition symbols*;
- Γ is a finite set of *observation symbols*;
- $O \subseteq \mathbb{R}$ is a finite set of *output symbols*;
- H is finite-dimensional Hilbert space called the *state space*;
- ρ is a map $\rho : \mathbb{R}_0^+ \to DO(H)$ which describes the evolution of the initial state of the automaton if no input arrives, such that $\rho(t)$ is a density operator in H, i.e. $\rho(t) \in DO(H)$, and $\rho(t)$ evolves according to the quantum postulates;
- $U = \{U_i\}_{i \in I}$ is a family of unitary transformations in H called the *family of transitions*;
- $A = \{A_\gamma\}_{\gamma \in \Gamma}$ is a family of observables over H called the *family of observables* such that the spectrum of A_γ is contained in O for all $\gamma \in \Gamma$.

The evolution of a QTPD with decoherence is the following. The initial state is evolving according to $\rho(t)$, and after the first input arrives at time t_0, the quantum automaton with decoherence behaves exactly as a common quantum automaton without decoherence with initial state $\rho(t_0)$. From this point on, the evolution of the automaton is ruled only by the unitary transitions and observations.

3. Transferring Zero-Knowledge Proofs

One of the main motivations for building quantum computers is that they are able to attack widely used classical cryptosystems, like RSA and ElGamal. Besides involved quantum algorithms, it is believed that simple quantum correlations may be used, in a distributed setting, to develop new attacks on security protocols. In this section, it is shown how QTPD's, together with Bell pairs, can be used to transfer proofs of an important class of zero-knowledge proof systems.

Here we show how QTPD's may be used to attack the impossibility of transferring the proof for the Goldreich–Micali–Wigderson system.[10] The importance of this system comes from the fact that all zero-knowledge proof systems can be reduced to ones similar to that of Goldreich, Micali and Wigderson,[10] and so the attack holds for all the protocols in this class.

A lot of research was pursued in understanding the power of zero-knowledge under computational restrictions; see, for instance, Ref. 11. For the attack presented

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

in this section it is assumed that all parties have the usual polynomial power (it could be quantum or probabilistic), and that the agents can use QTPD's.

A crucial feature of the attack is the use of tamper-proof Bell pairs. The tamper-proofing limits the verifier to measuring the first half of the Bell pair. Moreover, during the attack, the verifier can only perform two types of measurement over that half. These restrictions prevent the verifier from rolling back the memory state of the Bell pairs after performing measurements, allowing a third party to check if the verifier was interacting with a particular prover, as shown below.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

3.1. *Basic concepts*

Zero-knowledge proof systems have become one of the most important concepts in computational security since their invention by Goldwasser, Micali and Rackoff in 1985.[1] Before presenting zero-knowledge proof systems, we need to introduce several concepts from computational complexity. We assume that the reader knows the basics of computational complexity, namely the probabilistically oriented concepts (for a good reference, see Ref. 12).

To properly define an interactive proof system, it is necessary to consider a model allowing agents restricted to probabilistic polynomial time computations. There has been some discussion about the properties this model should satisfy and consequently about its expressiveness and behavior.[13–16] In this work, we introduce zero-knowledge proof systems in the same way as Goldwasser, Micali and Rackoff first introduced them in 1985.[1] However, contrary to what is common in this area, we will not make use of (interactive) Turing machines in proofs or when presenting algorithms, since that level of detail does not bring new insights as far as this paper is concerned. Instead we will refer to the Church–Markov–Turing postulate.[a] In the sequel, we denote the set $\{0,1\}$ by 2, and given a word $x \in 2^*$, we denote the *number of bits* in x (also called the *length* of x) by $|x|$.

Definition 3. An *interactive proof system* (V, P) is a protocol between two agents, the *verifier* V and the *prover* P, such that given an input $x \in 2^*$:

- Each agent has access to: (i) a private oracle that returns a (uniformly) random bit; (ii) the input x (only for reading); (iii) and a private memory.
- The protocol consists of several rounds and each agent is alternately active in each round. When active, each agent performs a computation and after that, either it sends a message to the other agent or the protocol halts.
- The computation of each agent depends only on the shared input x and on the messages received in the mean time from the other agent. The total time spent by the verifier V is polynomial in $|x|$.

[a]The Church–Markov–Turing postulate asserts that all the "reasonable" (classical) computational models are equivalent and interreducible in polynomial time (note that this reduction is not valid between classical and quantum models).

- When the protocol halts, V outputs a value [denoted by $(V,P)(x)$], and either accepts x [denoted by $(V,P)(x)_{\text{yes}}$] or rejects x.

In the following, given an interactive proof system (V,P), the probability of the verifier to accept the input x is denoted by $\text{Prob}((V,P)(x)_{\text{yes}})$. Note that what makes the computation of the verifier probabilistic is the fact that the oracle returns random bits. Moreover, observe that in an interactive proof system, only the verifier is restricted to doing polynomial time computations. There is no such restriction for the prover. So, for instance, given any NP language, it is not difficult to consider an interactive proof system in which the verifier accepts exactly the elements of that language, which we leave as an exercise to the reader.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

Example 2 (interactive proof system for SAT). The verifier V and the prover P share a propositional formula φ. The protocol runs as follows:

(1) If the formula φ is satisfiable, then P sends to V a valuation v that satisfies φ (note that the prover is not bounded to polynomial time computations and therefore it can find this v). If the formula is not satisfiable, then P sends a random valuation to V.
(2) When V receives a valuation v, it verifies whether the valuation satisfies φ. If this is the case it accepts φ; otherwise it rejects φ. In both cases, the output of the verifier, i.e. the value $(V,P)(\varphi)$, is the valuation v received from the prover.

Note that in Example 2, if $\varphi \in \text{SAT}$ then the verifier accepts φ with probability 1, and if $\varphi \notin \text{SAT}$ then the verifier accepts v with probability 0.

The formal definition of an interactive proof system is the following:

Definition 4. A *binary language* $L \subseteq 2^*$ *has an interactive proof system* if there exists an interactive proof system (V,P) such that:

(1) If $x \in L$ then $\text{Prob}((V,P)(x)_{\text{yes}}) > 2/3$.
(2) If $x \notin L$ then $\text{Prob}((V,P')(x)_{\text{yes}}) < 1/3$ for any prover P'.

The system (V,P) is said to be an *interactive proof system for* L.

One of the most significant results in the area of computational complexity concerning interactive proof systems was obtained by Shamir and states that PSPACE coincides with the class of languages for which there is an interactive proof system.[17] From a security point of view the concept of an interactive proof system is much more useful when enriched with the notion of zero knowledge. Informally, an interactive proof system is zero-knowledge if the verifier does not get any additional information by interacting with the prover. In the following, we will see how this additional property is important for obtaining certain security goals. Prior to defining what is zero-knowledge we need to introduce the concept of *computational indistinguishability*.

Definition 5. Two probabilistic algorithms A_1 and A_2 are *computationally indistinguishable* if, for any probabilistic polynomial time decision algorithm T, polynomial p and large-enough input x,

$$|\text{Prob}(T(A_1(x)) = 1) - \text{Prob}(T(A_2(x)) = 1)| \leq \frac{1}{p(|x|)}.$$

The notion of computational indistinguishability is important in cryptography for defining several security concepts, such as zero-knowledge proof systems.

Definition 6. *A* binary language $L \subseteq 2^*$ has a zero-knowledge proof system *if* L *has an interactive proof system* (V, P) *such that for any verifier* V' *there exists a probabilistic polynomial time algorithm* $S_{V'}$ *such that* (V', P) *is computationally indistinguishable from* $S_{V'}$.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

In what constitutes one of the fundamental results about zero-knowledge proof systems, it is shown in Ref. 10 that there is a zero-knowledge proof system for any language in NP provided that there is a nonuniform and indistinguishable cipher scheme. Zero-knowledge proof systems for languages in NP that are not believed to be in P can be employed to build useful security systems such as identification protocols,[2] e-voting protocols[3] and e-money protocols.[4] Informally theses systems allow the prover to show to the verifier that he knows a secret (a witness to an instance of the NP problem that is being considered) without revealing it. We now (re)define zero-knowledge proof systems from a cryptographic point of view.

Definition 7. Let L be a language in NP that is believed not to be in P, $x \in L$ and s a witness of x.[b] Let V (verifier) and P (prover) be two agents. P states that it knows a witness s of x, and V wants to check whether P really knows s. A zero-knowledge proof system for x and s is an interactive proof system (V, P) such that the following conditions hold:

- *Polynomial time bound* — The overall time complexity of both V and P is probabilistic polynomial in the size in bits of x.
- *Soundness* — If P knows a witness s, then V accepts that P has a witness s of x except with a negligible probability.[c]
- *Completeness* — If P does not know a witness s, then V accepts that P has a witness s of x with negligible probability.

[b]Recall that a language $L \subseteq 2^*$ is in NP if there exists a polynomial time decision algorithm A with two inputs, x and s, such that: (i) if $x \in L$ there exists a witness s such that $A(x, s) = 1$; (ii) if $x \notin L$ then for any witness s, $A(x, s) = 0$.

[c]A family of Bernoulli random variables $\{X_n\}_{n\in\mathbb{N}}$ is said to have negligible probability if, for any polynomial p and large-enough natural number n, $\text{Prob}(X_n = 1) < 1/p(n)$. In the context of soundness of zero-knowledge proof systems, the random variable X_n to consider takes value 1 if V does not accept that P has a witness s of x with size n.

- *Zero-knowledge* — During the protocol V does not get any new information about the secret witness s of x.
- *Impossibility of transferring the proof* — V cannot show to a third party that it was interacting with P.

Note that due to the polynomial time bound condition in Definition 7, an agent in a zero-knowledge proof system can only exchange a polynomial (in $|x|$) number of messages, and moreover the size of each message is also polynomially bounded over $|x|$. Furthermore, due to the zero-knowledge condition, a verifier does not obtain any new information about s, i.e. it must be possible to simulate, in a computationally indistinguishable way, the interaction between the prover and an arbitrary verifier using a probabilistic polynomial time simulator with input x.

It is worthwhile to briefly discuss and motivate each of the conditions presented in Definition 7. The polynomial time bound corresponds to the usual restriction in cryptography that all agents can only perform efficient computation. Soundness and completeness require that the verifier should accept that the prover has the secret (except with a negligible error) if and only if the prover actually knows the secret (i.e. a witness s of the instance x of the NP problem). The zero-knowledge property establishes that the verifier cannot acquire additional information about the secret s during the run of the protocol. In particular, if the goal of the proof system is to identify the prover, i.e. the secret s identifies the prover, the zero-knowledge property guarantees that the verifier cannot, after interacting with the prover, gain information to impersonate the prover. Finally, the impossibility of transferring the proof is an anonymity property. For instance, in identification protocols the impossibility of transferring the proof allows the anonymity of the prover in the sense that the verifier cannot show to a third party that a certain prover has identified to it. We will see that this property can be attacked when tamper-proof (quantum) memory is available.

One of the simplest examples of a zero-knowledge proof system according to Definition 7 is the Goldreich–Micali–Wigderson system,[10] which is based on the conjecture that deciding if two graphs are isomorphic is not in P although it is in NP.

Example 3 (Goldreich–Micali–Wigderson system). Suppose that both the verifier V and the prover P know two graphs G_0 and G_1 with n nodes. P says that it knows an isomorphism $\sigma : G_1 \to G_0$ (a witness to G_0 and G_1 being isomorphic). The protocol runs as follows:

(1) P generates a random isomorphism $\pi : G_0 \to H$ and sends H to V.
(2) V generates a random bit $b \in \{0, 1\}$ and sends it to P.
(3) P sends the isomorphism $\tau = \pi \circ \sigma^b$ to V (where $\sigma^0 = \mathrm{id}$).
(4) V checks if $\tau(G_b) = H$.

If the condition in Step 4 is not fulfilled, then V does not accept that P knows an isomorphism between G_0 and G_1, otherwise Steps 1–4 are repeated n times, and

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

the verifier only believes that P knows an isomorphism if in all these repetitions the condition in Step 4 holds.

It is interesting to observe that the problem chosen in Example 3, the graph isomorphism problem, is an NP problem that is not known to be NP-complete. In fact, it would be expected that NP-complete problems are in general more secure than general NP problems since there is a polynomial time reduction from all NP problems to NP-complete problems. Nevertheless, NP-complete problems have not been widely employed in cryptography as opposed to cryptographic systems based on problems such as the factorization and discrete logarithm (which belong to NP $\cap$ co-NP). The reason is that although NP-complete problems are very hard in the worst case, for some of them it is possible to find a witness for a random instance with high probability. Since the security of cryptographic systems relies on how hard it is to find those witnesses, NP problems believed to be sound to this attack, like graph isomorphism, are preferred.

Observe that the several steps of Example 3 constitute a typical iteration of a zero-knowledge proof system. The first step, designated by *commitment*, is when the prover sends the first message to the verifier. In the second step, called *challenge*, the verifier confronts the prover by typically sending it a bit. Finally, the third step, called the *decommitment*, is when the prover will have to send a message coherent with its commitment and with the challenge sent by the verifier. It can be shown that all the zero-knowledge proof systems for NP problems are reducible to a protocol with this structure.[10] The analysis and the attack presented in this section rely on this structure of the protocol (commitment, challenge and decommitment). The results can be straightforwardly generalized to any system with this structure although they are presented with respect to Example 3. It is interesting also to note that graph isomorphism seems not to be in BQP, i.e. it seems to be hard to attack graph isomorphism using quantum computers with the techniques known nowadays.[18]

The Goldreich–Micali–Wigderson zero-knowledge proof system described in Example 3 enjoys the following property as shown in Ref. 10.

Theorem 1 (Goldreich–Micali–Wigderson). *The protocol of Example* 3 *is a zero-knowledge proof system.*

The technique employed to prove that the Goldreich–Micali–Wigderson system is a zero-knowledge interactive proof system for classical adversaries has become a widely used method to show the security of several protocols[19] and a well-known technique to define security features.[13,15,16,20] This technique is called the *simulation paradigm*. It consists in showing that a real adversary of the protocol can be simulated by an ideal adversary, i.e. an adversary that cannot attack the system. For zero-knowledge proof systems the proof consists in showing that any dishonest verifier can be simulated by a process that does not interact with the prover (this process is the ideal verifier). Since the proof of Theorem 1 is important for understanding the results of the paper, we present it here.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

proof [Theorem 1]. Let V' be a verifier that might be dishonest, i.e. a verifier that does not follow the protocol specified in Example 3. The goal is to show that it is possible to simulate the interaction between this verifier V' and an honest prover P, without communication with the prover. First, observe that the purpose of the simulator is to generate n tuples (H, c, τ), where H is randomly generated (since the prover is honest) and c is a bit generated accordingly by V'. Observe as well that the simulator has access to the code of V'. In these conditions, a verifier V' is a family of probabilistic polynomial time algorithms $\{V'_k\}_{k\in 1\ldots n}$, where V'_k corresponds to the algorithm that V' uses to choose c in repetition k.

The algorithm V'_k receives a graph H and it may use some auxiliary data $w_k \in 2^*$ which it computed in previous repetitions. We assume that the initial data V' has about s is encoded in the string w_1. We now simulate V'_k as follows:

(1) The simulator chooses $b \in \{0, 1\}$ randomly (with uniform distribution).
(2) The simulator generates a random isomorphism $\tau : G_b \to H$ (with uniform distribution).
(3) The simulator applies $V'_k(H, w_k)$ and verifies if the bit c computed by V'_k is b:
    (a) If $c = b$ then the simulation of repetition k is successful, since the tuple generated, (H, c, τ), has the same distribution as one created by interacting with the prover as discussed below. The auxiliary computation done by V'_k, which can be used in the next repetition, is stored in w_{k+1};
    (b) If $c \neq b$ then the simulator jumps to Step 1.

Observe that the probability of $b = c$ is always $1/2$ (regardless of the computation of c by V'_k), since b was chosen uniformly. In other words, one expects that the simulator will have success in two trials. The simulator will not have success in n trials with probability $1/2^n$, which is negligible.

Finally, it is straightforward to show that the sequence of tuples (H, c, τ) generated by the simulator has precisely the same distribution of the tuples generated by an interaction between V' and P. For this reasons, these sequences are computationally indistinguishable. □

The simulation paradigm technique ensures the impossibility of transference of proof, i.e. the impossibility of tracing the interaction between the verifier and the prover. It is worthwhile to discuss this result since we will show how to attack it in a realistic context. The reason why the impossibility of transference of proof holds is the following. First, note that if the verifier wants to show to a third party — say, Eve — that he has been interacting with the prover, then he can send her the trace of the interaction, i.e. the sequence of tuples (H, c, τ). Nevertheless, as shown in Ref. 10, regardless of the behavior of the verifier, it is always possible to generate, without any communication with the prover, a sequence of tuples computationally indistinguishable from the trace of the interaction. So, if Eve has access to only the trace of the interaction, she may not be persuaded that the verifier did in fact interact with the prover, because that sequence of tuples may have been

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

generated by the verifier without communicating with the prover. For this point to hold the verifier must have complete control over the state of the computation. This scenario is too generous in what concerns the memory model of the verifier. If, at the beginning of the protocol, Eve gives a tamper-proof machine to the verifier such that he cannot put the machine in an arbitrary state, then it is not possible to apply the simulation paradigm. It is also interesting to note that if quantum memory is shared by Eve and the verifier it seems hard to apply the simulation paradigm. The reason for this is that if Eve and the verifier shared entangled memory and if measurements are performed by a dishonest verifier it seems hard to roll back. Therefore, the validity of the simulation paradigm in the quantum setting was an open problem. However, in Ref. 21, Watrous showed how to solve this problem. Again, in his proof Watrous assumes that the verifier has complete control over his private memory. As we will show in the next section in the context of the Goldreich–Micali–Wigderson system, if the memory of the verifier is tamper-proof, the verifier is able to show to Eve that he was interacting with the prover.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

3.2. *Quantum tamper-proof devices and transferring proofs*

We are now able to define the specific type of quantum tamper-proof machine that will perform the attack on the Goldreich–Micali–Wigderson system. Before that, we introduce the *computational observable* as $-|0\rangle\langle 0| + |1\rangle\langle 1|$ and the *diagonal observable* as $-|\nearrow\rangle\langle\nearrow| + |\searrow\rangle\langle\searrow|$ where $|\nearrow\rangle$ is $1/\sqrt{2}(|0\rangle + |1\rangle)$ and $|\searrow\rangle$ is $1/\sqrt{2}(|0\rangle - |1\rangle)$.

Definition 8. A *QTPD for a graph G with n nodes* is the quantum automaton $Q_G = (I, \Gamma, O, H, |\psi_0\rangle, U, A)$, where:

- $I = \emptyset$.
- $\Gamma = \{c, d\} \times \{1, 2\} \times \{1 \dots n\}^2$.
- $O = \{-1, 1\}$.
- $H = \otimes_{k=1}^{n} \otimes_{j=1}^{n} (\mathbb{C}^2 \otimes \mathbb{C}^2)$ (space constituted by n^2 pairs of *qubits*).
- $|\psi_0\rangle = \otimes_{k=1}^{n} \otimes_{j=1}^{n} (\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle))$.
- The family of observables A is such that:
  - $A_{(c,1,(k,j))}$ corresponds to the computational observable over the first *qubit* of the $(k \times j)$th pair with $k, j \in 1 \dots n$;
  - $A_{(d,1,(k,j))}$ corresponds to the diagonal observable over the first *qubit* of the $(k \times j)$th pair with $k, j \in 1 \dots n$;
  - $A_{(c,2,(k,j))}$ corresponds to the computational observable over the second *qubit* of the $(k \times j)$th pair with $k, j \in 1 \dots n$;
  - $A_{(d,2,(k,j))}$ corresponds to the diagonal observable over the second *qubit* of the $(k \times j)$th pair with $k, j \in 1 \dots n$.

  In general, $A_{(\underline{c},i,(k,j))}$ denotes the **c**omputational observable over the ith *qubit* of the $(k \times j)$th pair while $A_{(\underline{d},i,(k,j))}$ denotes the **d**iagonal observable over the same qubit.

To perform the attack the verifier has to show to Eve that he interacted with the prover after the interaction ended.[d]

In order to trace the interaction with a prover whose secret is an isomorphism between G_0 and G_1, Eve should set up a QTPD for G_0 with n nodes. This quantum automaton consists of n^2 pairs of qubits, all of them initially set at the state $1/\sqrt{2}(|00\rangle+|11\rangle) = 1/\sqrt{2}(|\nearrow\nearrow\rangle+|\searrow\searrow\rangle)$. The automaton accepts only observation inputs, and the observables are exactly the computational and diagonal observables for each qubit (so in total there are $2 \times 2 \times n^2$ observables). We index the n^2 Bell pairs by (k, j) with $k, j \in 1 \dots n$. To understand the notation of observation inputs, the symbol $(c, 1, (k, j))$ is associated with the **c**omputational observable of the **1**st half of the (k, j) Bell pair, and similarly $(d, 2, (k, j))$ is associated to the **d**iagonal observable of the **2**nd half of the (k, j) Bell pair.

We are now able to describe the protocol between Eve, the verifier and the honest prover, which will be used to obtain an evidence of the interaction between the verifier and the prover (this evidence is called a *trace of the interaction*). To this end, Eve and the verifier have to collude (nevertheless, notice that they do not trust each other, otherwise Eve would always believe in the verifier), in the sense that Eve will give the QTPD to the verifier, which will be used to store the trace. The tamper-proofing consists in the verifier being restricted to applying the observables $A_{(c,1,(k,j))}$ and $A_{(d,1,(k,j))}$.

Definition 9. *Protocol to trace the Goldreich–Micali–Wigderson system* for a prover knowing an isomorphism between G_0 and G_1 with n nodes.

(1) Eve and the verifier agree with a hash function

$$h : \mathcal{G}_n \to \{c, d\}^n,$$

where $\mathcal{G}_n$ is the set of all graphs with n nodes.

(2) Eve gives the sealed automaton Q_{G_0} to V. The automaton is sealed such that V can only execute the observation operators $A_{(c,1,(k,j))}$ and $A_{(d,1,(k,j))}$.

(3) V starts the interaction with P similarly to Example 3. For each repetition k:

   (a) In Step 1, V receives the graph H_k sent by P.

   (b) In Step 2, V computes $h(H_k) = e_{(k,1)} \dots e_{(k,n)}$ and applies the observables $A_{(e_{(k,j)},1,(k,j))}$ for all $j \in 1 \cdots n$ obtaining the sequence of measurements $o_{(k,1)} \dots o_{(k,n)}$ with $o_{(k,j)} \in \{-1, 1\}$. Next, V computes $b_k = \oplus_{j=1}^{n}(1 + o_{(k,j)})/2$ and sends $b_k \in \{0, 1\}$ to P, where $\oplus$ denotes the xor operator.

   (c) In Step 3, V receives the isomorphism τ_k sent from P [and checks if $\tau_k(G_{b_k}) = H_k$].

[d]Note that it is always assumed that the verifier cannot communicate with Eve during the interaction with the prover. Otherwise Eve would play the role in the verifier, and therefore could check herself that the prover is interacting. The attack consists in the verifier holding an evidence that he interacted with the prover after this interaction ended.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

(4) When the protocol ends, V sends the machine to Eve together with the sequence $w = (H_1, b_1) \cdots (H_n, b_n)$ of pairs (graph, bit) and the sequence $m = \tau_1 \cdots \tau_n$ of isomorphisms obtained in each repetition of the protocol.
(5) For each repetition $k = 1, \ldots n$, Eve will do as follows:

   (a) She starts by computing $h(H_k) = e_{(k,1)} \cdots e_{(k,n)}$ and applies the observables $A_{(e_{(k,j)},1,(k,j))}$ and $A_{(e_{(k,j)},2,(k,j))}$ for all $j \in 1 \cdots n$ obtaining the sequences $o_{(k,1)} \ldots o_{(k,n)}$ and $r_{(k,1)} \ldots r_{(k,n)}$, respectively, with $o_{(k,j)}, r_{(k,j)} \in \{-1, 1\}$.
   (b) Then, Eve checks if $o_{(k,j)} = r_{(k,j)}$ for all $j \in 1 \cdots n$ and if b_k (received from V) is equal to $\oplus_{j=1}^{n}(1 + r_{(k,j)})/2$.
   (c) Finally, Eve checks whether $\tau_k(G_{b_k}) = H_k$.

   If for all repetitions the checks performed in Steps 5(b) and 5(c) hold, then Eve believes that V was interacting with P.

The protocol above uses two features of quantum systems that assist in tracing the interaction between the verifier and the prover.

The first feature is the fact that the outcome of a measurement of a Bell pair is a uniform Bernoulli random variable. Therefore, the measurement performed by the verifier in Step 3(b) will provide a Bernoulli random variable for b_k, which Eve nor the verifier will be able to predict. Note that if no randomness was employed, and the sequence $b_1 \cdots b_n$ was predetermined by Eve, she could impersonate the prover while interacting from the verifier (jeopardizing completeness as presented in Definition 7). On the other hand, if the verifier has total control over the generation of the sequence $b_1 \cdots b_n$, there is no way Eve can distinguish a simulation performed solely by the verifier from the interaction between the verifier and the prover (since it is always possible to create efficiently a simulation for any sequence of challenges $b_1 \cdots b_n$). So, it is fundamental that the output of the machine is random, and that neither Eve nor the verifier has control over the sequence $b_1 \cdots b_n$. In order for the verifier to be sure that the machine states are Bell pairs, we can assume that a validity test is performed together with Eve when the QTPD is given to the verifier. In the light of this discussion, the QTPD must be tampered with for both Eve and the verifier. This explains why the only way Eve observes her half of the qubits is through the inputs of the QTPD in Step 4.

The second property of quantum systems that is used is entanglement. Bell pairs have the nice property that a measurement over one qubit will collapse the two qubits to the same state. Since Eve does not trust the verifier, she uses Bell pairs to be sure that the measurements performed by the verifier in one half of the pairs are consistent with the measurements she performs in Step 5(a) over the second half. We can assume the second half of the qubits are always in possession of Eve, and that when the machine is given to the verifier it contains only the first half of the Bell pairs.

Finally, it is possible to establish that the attack is possible.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

Theorem 2. *If the verifier can only change the state of the QTPD Q_{G_0} by executing $A_{(c,1,(k,j))}$ and $A_{(d,1,(k,j))}$, then the protocol presented in Definition 9 traces the interaction between the verifier and the prover for the Goldreich–Micali–Wigderson system.*

Proof. Assume that V wants to convince Eve that he was interacting with P without interacting with him.

By Step 4 of the protocol, V has to send to Eve a sequence $w = (H_1, b_1) \cdots (H_n, b_n)$. We split the repetitions into classes — those for which the verifier did not perform all measurements as described in Step 3(b) and those for which he does all the measurements.

Suppose that at repetition k the verifier did not perform in Q_{G_0} at least one measurement $A_{(e_{(k,j)},1,(k,j))}$ [as indicated in Step 3(b)] for $h(H_k) = e_{(k,1)} \cdots e_{(k,n)}$ and $k \in 1 \cdots n$. Then, this measurement will be performed by Eve during Step 5(a), and in that case the probability of $\oplus_{j=1}^{n}(1 + o_{(k,j)})/2$ being equal to b_k is $1/2$.

Thus, if there are m repetitions for which at least a measurement $A_{(e_{(k,j)},1,(k,j))}$ was not performed, the probability of the verifier cheating Eve and not being detected is at most $1/2^m$.

Now, we deal with the remaining $n-m$ repetitions. We will split these repetitions into two — those in which the verifier did not perform the measurements according to $h(H_k)$ and those where the measurements were made accordingly to $h(H_k)$. For the first case, it is straightforward to see that the probability of Eve detecting that V has cheated her, for each repetition, is $1/2$. Assuming that there are t of these repetitions, the probability of V not being detected is $1/2^t$. For the second case, we assume that there are $n-m-t$ repetitions. For each such repetition k, V obtains a sequence $o_{(k,1)} \cdots o_{(k,n)}$. Since $o_{(k,j)} \in \{-1, 1\}$ is a uniform Bernoulli random variable and $b_k = \oplus_{j=1}^{n}(1 + o_{(k,j)})/2$, we have $\text{Prob}(b_k = 1) = 1/2$. Considering all the $n-m-t$ graphs over all repetitions, the number of graphs N for which the verifier will not be able to construct the isomorphism efficiently (we assume that this problem is hard) is a random variable with binomial distribution with parameters $1/2$ and $n-m-t$. For the verifier not to be detected while cheating, he has to change the N graphs in the trace to send to Eve. Since the probability of finding collisions for h is negligible, we assume that the image by h of each replaced graph differs at least one bit from the image by h of the original graph. In these conditions, the probability of Eve not detecting that V is substituting the N graphs in Step 5 is $1/2^N$.

Combining the results of all repetitions, the probability of V not being detected is given by $1/2^{N+m+t}$. The expected value of $N+m+t$ is minimum when $m = t = 0$ and takes the value $n/2$, leading to probability $1/2^{n/2}$ which is negligible in n. □

4. Fair Contract Signing with QTPD's

In this paper, we show that quantum memory with a particular type of decoherence enables fairness in contract signing. Indeed, our assumptions are close to those used

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

by Crepeau[22] in order to achieve information-secure, quantum-oblivious transfer. We present only a simple solution for the problem, which, although fair and viable, has the disadvantage that an honest party who committed to the contract does not know whether the other party committed or not.

4.1. *Quantum contract signing*

The main idea of our contract signing protocol is to take advantage of decoherence in a tamper-proof device. In our particular case we consider the Werner states

$$\rho_{\eta(t)} = \eta(t)|\psi\rangle\langle\psi| + \frac{1-\eta(t)}{4} I, \tag{1}$$

decohering in time, where $|\psi\rangle$ is a maximally entangled state $|\psi\rangle = 1/\sqrt{2}(|00\rangle+|11\rangle)$ and $\eta(t)$ is a function specifying the decoherence in time. We use the function satisfying the property

$$|\eta(t) - 1| \leq \varepsilon \quad \text{for } t \leq t_0, \tag{2a}$$

$$|\eta(t) - 0| \leq \varepsilon \quad \text{for } t \geq t_1, \tag{2b}$$

for some fixed time points t_0 and t_1 (with $t_0 < t_1$) and a small nonnegative parameter ε. This time evolution describes a sudden decoherence given by the Werner state between time t_0 and t_1, and therefore any measurement performed on the maximally entangled state before t_0 will lead to perfect correlations, whereas measurements made after t_c will be completely uncorrelated. It is assumed that no agent has control over the Werner state decoherence, i.e. after time t_1 there is no entanglement between the parts of the Werner state, whatever the agents do. Moreover, notice that the partial trace of the Werner state is stable over time (and it is $(1/2)I$).

Under this memory model assumption it is possible to construct a fair contract signing protocol requiring no communication between Judge and the signing parties during the exchange phase (which is impossible in the classical case). To this end, assume that Alice and Bob each share with a trusted Judge $2n$ systems which are given by Werner states, as described in (1), and which completely decohere over the time period $[t_0, t_1]$. The objective is for Alice and Bob to commit to some contract at time t_0 and, upon committing at this time, it cannot be possible for one agent to decommit from the contract after time t_1. If any agent commits (or decommits) during the period $[t_0, t_1]$, then there is no guarantee whether the contract is binding or not. It is important to stress that whatever an agent does, in any period of time, he cannot make the contract both binding and unbinding without negligible probability. The solution is given in Protocol 1.

During the certification procedure, if one agent — say, Alice — wants to enforce the contract, she asks Judge to bind the protocol. Upon this request Judge compares the part of the Werner qubits shared with Alice and measured in the computational basis obtaining a sequence j_A. Then Judge asks Alice to show the result r_A she obtained when she measured her qubits. If r_A and j_A coincide, Judge gives the right to Alice to check on Bob's commitment to the contract.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

For the protocol not to be binding, Bob has to show that he did not commit to the contract. Bob can do this by showing to Judge the result r_B of measuring with the diagonal basis his part of the Werner states shared with Judge. Judge compares the outcomes of measurements performed on his systems in the diagonal basis obtaining j_B. If j_B and r_B coincide, then the contract is not binding; otherwise it is binding. Note that if Bob did measure his part of the Werner states with the computational basis (i.e. he committed to the contract), then he will not be able to produce the sequence j_B unless with an exponentially low probability. The proof of this result follows:

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

Protocol 1: Fair contract signing without bindingness verification

Participants: Alice, Bob, trusted Judge.
Input: Times t_0 and t_1 (with $t_0 < t_1$), security parameter ε.
Goal: Alice and Bob want to sign a document simultaneously.
Requirement: Alice and Bob each share with Judge $2n$ systems given by Werner states that decohere as specified by Eq. (1).

1: Alice signs the document and sends it to Bob (before time t_0).
2: Bob signs the document and sends it to Alice (before time t_0).
3: If Alice wants to commit to the contract, she measures her $2n$ qubits before time t_0 in the computational basis; otherwise she measures the $2n$ qubits in the diagonal basis.
4: If Bob wants to commit to the contract, he measures his $2n$ qubits before time t_0 in the computational basis; otherwise he measures the $2n$ qubits in the diagonal basis.
5: Judge chooses at random n of his qubits (shared with Alice) to measure in the computational basis and he measures the other n qubits (shared with Alice) in the diagonal basis. He measures his qubits whenever he wants. He does the same procedure for the $2n$ qubits shared with Bob.
6: If later Alice (or Bob) wants to enforce the contract, she sends the outcomes of her measurement to Judge, who challenges Bob to state whether he committed or rejected the contract and to support his statement with the measurement outcomes.
7: Judge compares his measurements in the corresponding basis with Alice's and Bob's outcomes.
8: **if** all measurement outcomes are consistent **then**
9: If both decisions are *accept*, then Judge accepts; otherwise he rejects.
10: **else if** measurement outcomes of either Alice or Bob are inconsistent **then**
11: Judge enforces decision of the consistent participant.
12: **else** ▷ *both Alice and Bob supplied inconsistent measurement outcomes*
13: In this case there is no honest participant to protect.
14: **end if**

Theorem 3. (1) *If both agents committed to a contract before time* t_0 *using* n *Werner states, then the probability of Judge unbinding the contract is exponentially low in the size of* n*, whatever the agents did before time* t_1 (binding soundness). *Similarly,* (2) *if one agent unbound the contract before time* t_0 *using* n *Werner states, then Judge will bind the contract after time* t_1 *with an exponentially low probability in the size of* n (unbinding soundness). *Finally,* (3) *there is no strategy that an agent can perform that will make the contract both binding and unbinding without negligible probability in the size of* n*, after time* t_1 (completeness).

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

Proof. The binding soundness is easy. If both Alice and Bob committed to the contract before time t_0, i.e. measured their part of systems in the computational basis, Judge's qubits and the agents' qubits are not entangled after time t_0. Moreover, Judge's qubits have collapsed to the computational basis. In this case, the probability that an agent is able to give the sequence of diagonal measurements for Judge's qubits is 2^{-n}. The unbinding soundness is similar.

Completeness follows by noticing that there is no entanglement between Judge's system and the agent's system after time t_1. We will now show by *reductio ad absurdum* that there cannot be a strategy by Alice/Bob that will make the contract both binding and unbinding with nonnegligible probability. The separation condition requires that Alice/Bob should not affect Judge's system after t_1. This means that if Alice/Bob can attack the protocol, they must have a prediction of Judge's measurements for both the diagonal and the computational basis *before* time t_1. So assume that — say, Alice — has a family of POVM's $E^n = \{E^i\}_{i \in M_n}$ (where $2n$ is the number of shared qubits with Judge and M_n is the index family denoting the outcomes of the POVM) that will allow her to generate two bit strings c and d of size n that will predict, before time t_1 and with *nonnegligible probability in* n, the outputs of Judge's qubits in both the computational and the diagonal basis. We will now show that such a family of POVM's does not exist, even in the case of a joint POVM performed on all $2n$ qubits.

Assume that the way Judge chooses his measurement basis is uniform and independent of any computation that Alice does. This means that the best strategy that Alice can do if she predicts for qubit i — say $c_i = |0\rangle$ and $d_i| \nearrow\rangle$ — is to, using a suitable POVM, try to prepare the ith qubit of Judge to

$$\cos\left(\frac{\pi}{8}\right)|0\rangle + \sin\left(\frac{\pi}{8}\right)|1\rangle, \tag{3}$$

thus maximizing the probability of success of achieving either 0 or + [this value is $\cos^2(\pi/8)$]. In this way, the probability of Alice being detected is $\sin^2(\pi/8)$. This value is a lower bound of the detection probability that the family of POVM's $E^n = \{E^i\}_{i \in M_n}$ can do, since it is not guaranteed that the POVM will be able to set Judge's qubit in (3) (for $c_i = |0\rangle$ and $d_i| \nearrow\rangle$). It is easy to see that for all possible combinations of c_i and d_i, the best strategy for Alice is to set Judge's qubit in some state similar to that in (3), which will lead to detection with probability

$\sin^2(\pi/8)$. Therefore, even in the best scenario, the probability of Judge detecting that Alice is tricking is $\sin^{2n}(\pi/8)$, which is negligible in n. □

4.2. *Noise and being tamper-proof*

It is important that the noise that transfers $\rho_{\eta(0)}$ to $\rho_{\eta(t_1)}$ [see Eq. (1)] affects the systems controlled by Alice and Bob, because otherwise all actions of Alice and Bob (namely the measurement) commute with the noise. On the other hand, it is sufficient if the noise affects only the systems controlled by Alice and Bob, i.e. the noise at Alice and Bob has to be tamper-proof. There is no need for the decoherence on the quantum systems controlled by Judge. To be more formal, the quantum automaton with decoherence that we need for implementing the protocol is $Q = (I, \Gamma, O, H, \rho, U, A)$, where:

- $I = \emptyset$;
- $\Gamma = \{c, d\}$;
- $O = \{0, 1\}^{2n}$ is the set of outcomes of the measurements;
- $H = (\mathbb{C}^2 \otimes \mathbb{C}^2)^{2n}$ is the state space required for $2n$ maximally entangled pairs, where the first qubit is in the possession of Alice and the other with Judge;
- ρ is given by Eq (1);
- A_c corresponds to the observable associated to measuring all Alice's qubits in the computational basis;
- A_d corresponds to the observable associated to measuring all Alice's qubits in the diagonal basis.

We can achieve the same goal as in Protocol 1 using only single-qubit states. Judge can create $4n$ qubits randomly either in the standard or the dual basis state and send them to Alice and Bob. Afterward — to check their commitment — he uses the knowledge of the distributed states instead of the measurement.

4.3. *Analysis of the protocol*

Unfortunately, Protocol 1 has one very unpleasant property. It is fair in the sense that no participant X can achieve the situation where he is sure that the other participant Y bound the signature and at the same time Y is not sure whether X bound the protocol. Also, the protocol is correct in the sense that if both participants want to sign the document and they behave honestly, the document is signed. It also true that in the contract exchange phase Judge does not need to be contacted. However, without contacting Judge, the participant can never be sure whether the other agent's signature is binding or not, which is an unpleasant property in many situations. It is possible to improve this situation by including a protocol that increases the confidence of Alice and Bob in the commitment of each other to the contract. Improvements are explored in Ref. 23.

On the other hand, we may easily find situations (cheating models) where Protocol 1 is perfectly sufficient. We specify them in terms of game theory and give

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

some examples. Let S_A, S_B, R_A, R_B denote Alice and Bob's decision to sign/reject. We use the payoff functions to specify individual preferences of Alice and Bob:

	Payoff for Alice		Payoff for Bob	
	S_A	R_A	S_A	R_A
S_B	w_A	x_A	w_B	x_B
R_B	y_A	z_A	y_B	z_B

The values w, x, y, z denote respective payoffs for Alice and Bob. Our protocol works perfectly as long as $w > x, y, z$ for both Alice and Bob, i.e. the most preferred situation for both participants is that the contract is signed. An example is an online plane ticket purchase — the travel agent wants to be sure that he can cash the received electronic cheque, and in case he cannot, the customer cannot use the ticket. The other way around, the customer wants to be sure that he receives the ticket before the travel agent cashes the money. Most of them do not have the objectives of tricking the other party by rejecting the contract and making the other party believe that it is signed — this helps only to make financial losses for the travel agent and tricking the customer with an invalid ticket.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

5. Conclusion

The use of classical tamper-proof devices to attack classical protocols is further investigated in Ref. 24. Observe that a QTPD can almost be simulated by a classical tamper-proof machine, but several problems arise. Randomness has to be replaced by pseudorandomness, and moreover decoherence has to be replaced by a global clock, which is not at all straightforward to achieve.

Concerning the attack on the impossibility of transferring proofs of the Goldreich–Micali–Wigderson system, it is more or less obvious that one should adapt the attack to zero-knowledge proof systems based on the commitment, challenge and response structure, which account for the majority of the zero-knowledge proof systems (an exception is noninteractive zero-knowledge systems that require a common random string). Attacks using only classical tamper-proof devices are developed in Ref. 24, namely attacks on undeniable and ring signatures.

Work is also being carried out on improving the contract singing protocol. The idea is to include a protocol that increases the confidence of Alice and Bob in the commitment of each other to the contract. The improvements are explored in Ref. 23.

Acknowledgments

This work was partially supported by FCT (Portugal) and EU FEDER through projects QSec PTDC/EIA/67661/2006 and QuantLog POCI/MAT/55796/2004. N. Paunkovic is grateful for the support from FCT (Portugal) and EU FEDER grant SFRH/BPD/31807/2006. J. Bouda acknowledges the support from grants

GAČR 201/07/0603, MŠM0021622419 and GAČR 201/06/P338. We also thank the anonymous referees whose suggestions significantly improved the paper.

References

1. S. Goldwasser, S. Micali and C. Rackoff, The knowledge complexity of interactive proof systems, *SIAM J. Comput.* **18** (1989) 186–208.
2. A. Fiat and A. Shamir, How to prove yourself: Practical solutions to identification and signature problems, in *CRYPTO' 86*, ed. A. M. Odlyzko, Lecture Notes in Computer Science, Vol. 263 (Springer, 1987), pp. 186–194.
3. R. Cramer, R. Gennaro and B. Schoenmakers, A secure and optimally efficient multi-authority election scheme, in *EuroCrypt' 97*, ed. W. Fumy, Lecture Notes in Computer Science, Vol. 1233 (Springer, 1997), pp. 103–118.
4. D. Chaum, A. Fiat and M. Naor, Untraceable electronic cash, in *Crypto '88*, ed. S. Goldwasser, Lecture Notes in Computer Science, Vol. 403 (Springer, 1990), pp. 319–327.
5. S. Even and Y. Yacobi, Relations among public key signature systems, *Technical Report 175* (Technion, 1980).
6. M. J. Fischer, N. A. Lynch and M. Paterson, Impossibility of distributed consensus with one faulty process, *J. ACM* **32** (1985) 374–382.
7. M. Ben-Or, O. Goldreich, S. Micali and R. L. Rivest, A fair protocol for signing contracts, *IEEE Trans. Inform. Theory* **36** (1990) 40–46.
8. N. Asokan, M. Schunter and M. Waidner, Optimistic protocols for fair exchange, *ACM Conf. Comput. Commun. Security* (1997), pp. 7–17.
9. C. Moore and J. P. Crutchfield, Quantum automata and quantum grammars, *Theoret. Comput. Sci.* **206** (2000) 275–306.
10. O. Goldreich, S. Micali and A. Wigderson, Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems, *J. ACM* **38** (1991) 690–728.
11. C. Dwork and L. Stockmeyer, Finite state verifiers I: The power of interaction, *J. ACM* **39** (1992) 800–828.
12. C. H. Papadimitriou, *Complexity Theory* (Addison-Wesley, 1994).
13. B. Pfitzmann and M. Waidner, Composition and integrity preservation of secure reactive systems, *CCS '00* (2000), pp. 245–254.
14. R. Canetti, I. Damagrd, S. Dziembowski, Y. Ishai and T. Malkin, On adaptive versus non-adaptive security of multiparty protocols, in *EUROCRYPT' 01* (Springer, 2001), pp. 262–279.
15. J. Mitchell, A. Ramanathan, A. Scedrov and V. Teague, A probabilistic polynomial-time calculus for analysis of cryptographic protocols, to appear in *Theoret. Comput. Sci.*
16. P. Mateus, J. Mitchell and A. Scedrov, Composition of cryptographic protocols in a probabilistic polynomial-time process calculus., in *CONCUR '03*, eds. R. Amadio and D. Lugiez, Lecture Notes in Computer Science, Vol. 2761 (Springer-Verlag, 2003), pp. 327–349.
17. A. Shamir, IP = PSPACE, *J. ACM* **39** (1992) 869–877.
18. C. Moore, A. Russell and J. Schulman, The symmetric group defies strong Fourier sampling, *FOCS '05* (2005), pp. 479–490.
19. S. Goldwasser, Multi-party computations: Past and present, in *PODC '97* (ACM, 1997), pp. 1–6.
20. R. Canetti, Security and composition of multi-party cryptographic protocols, *J. Cryptol.* **13** (2000) 143–202.

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

21. J. Watrous, Zero-knowledge against quantum attacks, in *STOC '06* (ACM, 2006), pp. 296–305.
22. C. Crépeau, Quantum oblivious transfer, *J. Mod. Opt.* **41** (1994) 2445–2454.
23. J. Bouda, P. Mateus and N. Paunkovic, Quantum contract signing, *Technical Report SQIG-IT*, submitted for publication (2008).
24. P. Mateus and S. Vaudenay, Defeating privacy enhancement tools with seals, *Technical Report SQIG-IT*, submitted for publication (2008).

Int. J. Quantum Inform. 2008.06:281-302. Downloaded from www.worldscientific.com
by EAST CAROLINA UNIVERSITY on 04/19/15. For personal use only.

International Journal of Quantum Information
Vol. 4, No. 3 (2006) 573–583
© World Scientific Publishing Company

World Scientific
www.worldscientific.com

DISCRETE TIME QUANTUM WALK ON A LINE WITH TWO PARTICLES

L. SHERIDAN
Institute for Quantum Computing, University of Waterloo,
200 University Avenue West, Waterloo,
Ontario N2L 3G5, Canada
lssherid@iqc.ca

N. PAUNKOVIĆ
SQIG, Instituto de Telecommunicações and Instituto Superior Técnico,
P-1049-001 Lisbon, Portugal
paunkovic@isiosf.isi.it

Y. OMAR
CEMAPRE ISEG, Universidade Técnica de Lisboa,
P-1200-781 Lisbon, Portugal
yasser.omar@iseg.utl.pt

S. BOSE
Department of Physics and Astronomy, University College London,
Gower Street, London WC1E 6BT, UK
sougato.bose@qubit.org

Received 30 May 2005

We introduce the idea of a quantum walk with two particles and study it for the case of a discrete time walk on a line. We consider both separable and maximally entangled initial conditions, and show how the entanglement and the relative phase between the states describing the *coin* degree of freedom of each particle will influence the evolution of the quantum walk. In particular, these factors will have consequences for the distance between the particles and the probability of finding them at a given point, yielding results that cannot be obtained from a separable initial state, be it pure or mixed. Finally, we review briefly proposals for implementations.

Keywords: Quantum walks; entanglement; quantum information.

1. Introduction

Quantum walks, first proposed in 1993,[1] are the quantum analog of classical random walks. Quantum walks have recently generated much interest (for an overview, see Ref. 2), in particular because they are proving to be a very useful technique for the construction of quantum algorithms, just as random walks are for classical

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

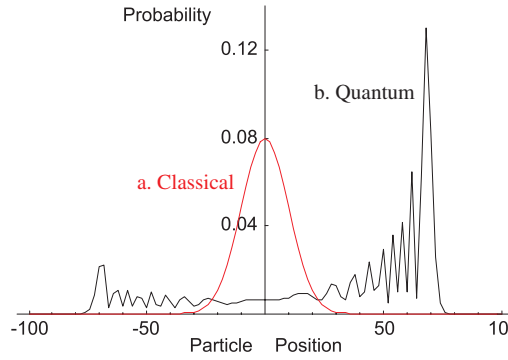


Fig. 1. Probability distributions for discrete time random walks on a line after $N = 100$ steps: (a) the classical walk and (b) the quantum walk with a Hadamard coin and initial state $|0\rangle \otimes |1\rangle$.

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

algorithms. In fact, several quantum algorithms based on quantum walks have been shown to be optimal,[3,4] and the only algorithm currently proven to offer an exponential speed-up with respect to its classical counterpart[5] is based on a (continuous time) quantum walk.

The crucial difference between quantum and random walks is that the former allow for quantum superpositions of the walker states and explore the interference of the terms in these superpositions. The resulting probability distributions after N steps are very different, as can be seen in Fig. 1. Furthermore, a quantum walk exhibits a variance proportional to N, which represents a quadratic speed-up over the classical case, where the variance goes as $\sqrt{N}$.

2. The Multiparticle Walk

The new concept introduced in this paper is the idea of using multiple walkers — quantum particles — simultaneously in the same quantum walk.[6] This idea has been used classically, where k walkers traversing the same graph are equivalent to k independent random walks. These additional resources can be used, for instance, to improve the time to hit a certain node of the graph. In the quantum case though, the multiple-walker situation becomes even more interesting, because the particles can be entangled, which results in final probability distributions which cannot be described by multiple independent single-particle walks. We will show that the entanglement in a quantum walk on a line can be tuned to cover more space.

3. The Quantum Walk Formalism

Let us now describe a single-particle discrete time quantum walk on a line. A single step of the walk comprises two operations: the coin flip $\hat{C}$ and the conditional shift $\hat{S}$. The coin degree of freedom is encoded by a qubit, whose states we will denote $|\uparrow\rangle$ and $|\downarrow\rangle$, and the coin flip is a unitary operator that sets a basis state into a superposition over that basis. We have that $\hat{C} \in SU(2)$ acts on the coin Hilbert

space $\mathcal{H}_C$. The most commonly used example is the Hadamard transformation, and we shall use this throughout as the coin:

$$\hat{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{1}$$

This coin is said to be *unbiased* in the sense that an initial state $|\uparrow\rangle$ will be set into an equal superposition of "up" and "down," $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$, after being acted upon by the coin. This corresponds to equal probabilities of the particle being found in the state $|\uparrow\rangle$ or $|\downarrow\rangle$ if measured after one operation. The second part of a single step of the walk is the shift operation $\hat{S}$, acting on $\mathcal{H}_P$ — the Hilbert space encoding the possible positions of a particle on an infinite discrete line. The basis vectors of $\mathcal{H}_P$ will be denoted $|i\rangle$, where $i \in \mathbb{Z}$. Thus, the total Hilbert space of the walk is given by

$$\mathcal{H} = \mathcal{H}_C \otimes \mathcal{H}_P, \tag{2}$$

and the states of the quantum walk will be described by vectors in $\mathcal{H}$.

We will consider the following shift operator:

$$\hat{S} = \sum_i \left(|\uparrow\rangle \langle\uparrow| \otimes |i+1\rangle \langle i| + |\downarrow\rangle \langle\downarrow| \otimes |i-1\rangle \langle i| \right). \tag{3}$$

This has the effect of moving a particle in the "up" state one unit to the right and a particle in the "down" state one unit to the left. If the particle in question is in some superposition of "up" and "down," the terms evolve accordingly. $\hat{C}$ acts on the coin state space and $\hat{S}$ acts on both the position state space and coin space together. The total operator for each step is unitary and will have the form:

$$\hat{U} = \hat{S}(\hat{I}_P \otimes \hat{C}), \tag{4}$$

where I_P is the identity operator on the position space. So, for example, if the initial state is $|\uparrow\rangle \otimes |0\rangle$, the first step of the walk gives

$$\hat{U}(|\uparrow\rangle \otimes |0\rangle) = \frac{1}{\sqrt{2}}(|\uparrow\rangle \otimes |1\rangle + |\downarrow\rangle \otimes |-1\rangle). \tag{5}$$

If a measurement is performed at this point, then the walk agrees with its classical counterpart: there is a 50% probability of the particle being found at position $+1$ and a 50% probability of the particle being found at position -1. Yet, after the first two steps of the walk, the progression of the quantum and classical walks begin to diverge wildly. This becomes particularly clear after a few dozen steps. In Fig. 1, we present the probability distributions for both the classical random walk and the quantum walk for $N = 100$.

Note that the quantum walk shows a relatively low probability associated with the walker being found close to the origin: rather, the peaks of the distribution correspond to the particle being found a considerable distance away. On the other hand, in the classical case, the origin is precisely the point where the probability for the particle to be found is maximal. There is a $\sqrt{N}$ speed-up of the quantum walk over the classical in terms of expected deviation of the walker from the origin.

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

4. Quantum Walk with Two Particles

Let us now generalize this definition of the discrete time quantum walk on a line to the case of two walkers. Consider two particles, 1 and 2, simultaneously completing quantum walks on the same line. Let the joint Hilbert space of the two-particles system be $\mathcal{H}_{12}$. Then, we have

$$\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2, \tag{6}$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ are the Hilbert spaces of particles 1 and 2, respectively. Both $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic to $\mathcal{H}$, as defined in Eq. (2). Similarly, the new walk operator will be $\hat{U}_{12}$, given by

$$\hat{U}_{12} = \hat{U} \otimes \hat{U}. \tag{7}$$

Now, suppose that two particles are set into the walk in a pure separable initial state, for example:

$$|\Psi_0^{\text{sep}}\rangle = |0,\downarrow\rangle_1 \, |0,\uparrow\rangle_2 \, . \tag{8}$$

In this case, the two particles evolve throughout the walk simultaneously, but independently, so that after N steps:

$$|\Psi_N^{\text{sep}}\rangle = \hat{U}_{12}^N \, |\Psi_0^{\text{sep}}\rangle = \hat{U}^N \, |0,\downarrow\rangle_1 \, \hat{U}^N \, |0,\uparrow\rangle_2 \, . \tag{9}$$

This is equivalent to two separate walks being completed on two different lines at the same time. The probability distributions are both identical (though one is inverted in the position coordinate relative to the other) and independent.

However, this is a quantum mechanical scheme, so it is also possible for a pair of particles to be set into the walk in a joint state which is entangled. For example, consider two maximally entangled states:

$$|\psi_0^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|0,\downarrow\rangle_1|0,\uparrow\rangle_2 \pm |0,\uparrow\rangle_1|0,\downarrow\rangle_2). \tag{10}$$

Note that these also describe the cases of a pair of identical particles on the same point, either bosons (the "+" state) or fermions (the "−" state). Now the evolution of the walk cannot be described as two separate walks progressing independently. The joint state of the particles' evolution after N steps is

$$|\psi_N^{\pm}\rangle_{12} = \hat{U}_{12}^N \, |\psi_0^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}\big(\hat{U}^N \, |0,\downarrow\rangle_1 \; \hat{U}^N \, |0,\uparrow\rangle_2 \pm \hat{U}^N |0,\uparrow\rangle_1 \, \hat{U}^N |0,\downarrow\rangle_2\big). \tag{11}$$

These two scenarios correspond to very different final joint probability distributions. Let $P_{12}(i,j,N)$ denote a joint probability over the two particles of finding particle 1 in position i and particle 2 in position j after N steps of the walk. In the case of the separable initial conditions, $|\Psi_0^{\text{sep}}\rangle = |0,\downarrow\rangle_1 \, |0,\uparrow\rangle_2$,

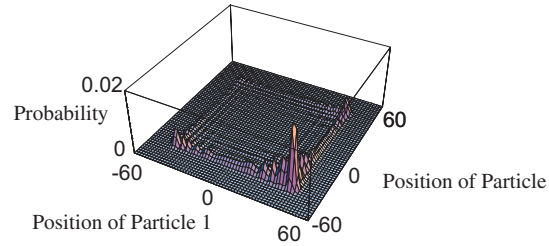
$$P_{12}^{\text{sep}}(i,j;N) = P_1^{\text{sep}}(i;N) \times P_2^{\text{sep}}(j;N), \tag{12}$$

which is just the product of two independent distributions for single-particle walks after N steps. This is in agreement with the observation that the two particles

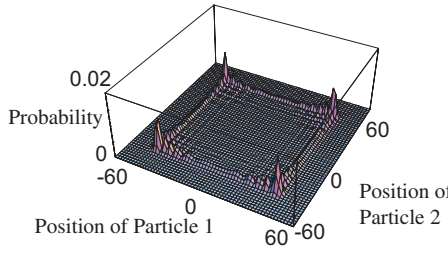
in the walks proceed independently, and without reference to each other, into the final state given in Eq. (9). However, this equation does not describe the case of entangled particles, which have walk evolutions which depend on the existence of the other particle in the walk. The joint probability distribution P_{12}^{pm} must be symmetric under exchange of the labels 1 and 2 for the "+" case and, for the "−" case, the distribution must remain unchanged when *both* the labels of the particles are exchanged and the position axes are reflected. These distributions are presented graphically in Fig. 2 for the case $N = 60$. Looking at the plots, it is clear that these distributions are significantly different.

In the case of the separable initial conditions, the plot in Fig. 2(a) is just the product of two distributions like the one in Fig. 1(b). The initial state is $|\Psi_0^{\text{sep}}\rangle = |0,\downarrow\rangle_1 \, |0,\uparrow\rangle_2$, so the distribution for particle 2 is biased towards the right [as shown in Fig. 1(b)] and the distribution for particle 1 is biased towards the left — a mirror image of the distribution for particle 2. Figure 2(b) shows the distribution for P_{12}^{+} and Fig. 2(c) shows the distribution for P_{12}^{-}.

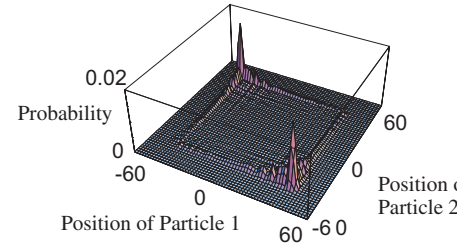
In all three distributions, the maxima occur around $i, j \simeq 42$, but the effect of the entanglement is dramatic. Entanglement can induce the walk to explore certain configurations with relatively high probability that in a walk beginning in a separable state would be very unlikely. For example, the higher maxima of the distribution for the symmetric entangled state ("+" state) correspond to the particles being found on the same side of the origin and both near $|i|, |j| \simeq 42$.



(a) Separable initial conditions.



(b) Entangled initial conditions "+".



(c) Entangled initial conditions "−".

Fig. 2. Two-particle probability distributions after $N = 60$ steps for different initial conditions: (a) separable state $|\psi_0^S\rangle_{12}$, (b) $|\psi_0^+\rangle_{12}$ state, and (c) $|\psi_0^-\rangle_{12}$ state. Note that the peak of the distribution in (a) has been cropped.

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

Conversely, in the antisymmetric "−" case, note that a pair of identical fermions conducting a quantum walk will follow the distribution shown in Fig. 2(c), in which the particles have high probability of being found at opposite ends of the line. A pair of identical bosons will follow the symmetric distribution shown in Fig. 2(b).

The behaviors of the individual particles in the quantum walk for entangled initial conditions also exhibit interesting features. Tracing over the degrees of freedom of one of the particles leaves a reduced density matrix:

$$\rho_1 = \mathrm{Tr}_2\left(|\psi_N^{\pm}\rangle_{12}\langle\psi_N^{\pm}|_{12}\right) = \frac{1}{2}\hat{U}^N|0,\downarrow\rangle\langle 0,\downarrow|\hat{U}^{\dagger N} + \frac{1}{2}\hat{U}^N|0,\uparrow\rangle\langle 0,\uparrow|\hat{U}^{\dagger N}, \quad (13)$$

which is an equal mixture or the states $U^N|0,\downarrow\rangle$ and $U^N|0,\uparrow\rangle$. These are simply the evolution states after N steps for a walk starting from the state $|0,\downarrow\rangle$ and the state $|0,\uparrow\rangle$. Therefore, the marginal probability distribution for finding one of the particles in position i after N steps is given by

$$P_1^{\pm}(i;N) = \frac{1}{2}[P_{\downarrow}(i;N) + P_{\uparrow}(i;N)] = P_2^{\pm}(i;N). \quad (14)$$

Note that $P_{\downarrow}(i;N)$ and $P_{\uparrow}(i;N)$ are the probability distributions of a single-particle walk initialized in the state $|0,\downarrow\rangle$ and $|0,\uparrow\rangle$. Again compare this situation to the case of separable states. The separable initial condition given in Eq. (8) generates the marginal probability distributions $P_1^{\mathrm{sep}}(i;N) = P_{\downarrow}(i;N)$ and $P_2^{\mathrm{sep}}(i;N) = P_{\uparrow}(i;N)$, yielding a joint probability distribution which is just the product of these [Eq. (12)]. This is not the case for the entangled state, where the joint probability distribution contains information about the correlations between the positions of the two particles, which will become apparent on measurement of the two particles' positions.

There are further joint properties of the two particles which are of interest. Let us define the distance $\Delta_{12}^{\mathrm{sep},\pm}$ between the two particles' final (independently) measured positions, x_1 and x_2, after N steps:

$$\Delta_{12}^{\mathrm{sep},\pm} \equiv |x_1 - x_2|, \quad (15)$$

where $x_1, x_2 \in [-N,\ldots,0,\ldots,N]$. Table 1 presents the expectation value of this distance for the three different initial conditions, and for different N. From this data, it is evident that the particles are more likely to remain closer together for the "+"-entangled initial conditions than for the separable initial state, and more likely to remain farther apart in the case of the "−"-entangled initial state than for the separable state. In fact, we can make a stronger statement than this; for a given N, we always have

$$\langle\Delta_{12}^{-}\rangle - \langle\Delta_{12}^{\mathrm{sep}}\rangle = \langle\Delta_{12}^{\mathrm{sep}}\rangle - \langle\Delta_{12}^{+}\rangle. \quad (16)$$

Let us now consider the correlation function between the spatial distribution of each of the two particles:

$$C^{\mathrm{sep},\pm}(x_1,x_2) \equiv \langle x_1 x_2\rangle - \langle x_1\rangle\langle x_2\rangle. \quad (17)$$

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

Table 1. Expectation value $\langle \Delta_{12}^{\text{sep},\pm} \rangle$ after N steps.

Number of steps N	10	20	30	40	60	100
Initial condition $\|\psi_0^-\rangle_{12}$	8.8	17.5	26.0	34.9	52.2	87.0
Initial condition $\|\psi_0^{\text{sep}}\rangle_{12}$	7.1	14.7	21.9	29.5	44.3	73.9
Initial condition $\|\psi_0^+\rangle_{12}$	5.5	11.9	17.8	24.1	36.3	60.8

Table 2. Correlation function $C^{\text{sep},MS,\pm}(x_1, x_2)$ after N steps.

Number of steps N	10	20	30	40	60	100
Initial condition $\|\psi_0^-\rangle_{12}$	-16.8	-69.8	-153.5	-276.2	-619.7	-1718.3
Initial condition $\hat{\rho}_{12}^{MS}(0)$	-6.0	-31.4	-70.6	-130.4	-299.3	-839.3
Initial condition $\|\psi_0^{\text{sep}}\rangle_{12}$	0	0	0	0	0	0
Initial condition $\|\psi_0^+\rangle_{12}$	4.8	7.3	13.7	15.1	23.1	39.1

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

Clearly, in the case of the separable initial condition (8), this correlation is always zero. For the other cases, the values of $C^{\pm}(x_1, x_2)$ are presented in Table 2 for different N. Given the symmetry of the $P_{12}^{\pm}$ distributions, we have that in those cases $\langle x_1 \rangle = \langle x_2 \rangle = 0$. Thus, the sign difference in the correlation function expresses the tendency for the two particles in the "$-$" case to end the quantum walk on different sides of the line (with respect to the origin, 0), and on the same side for the "$+$" case.

Finally, let us now calculate, for the different initial conditions, the probability of finding at least one particle in position i after N steps: $\mathcal{P}^{\text{sep},\pm}(i; N)$. This is clearly a joint property as it depends on both one-particle outcomes:

$$\begin{aligned}\mathcal{P}^{\text{sep},\pm}(i; N) &= \sum_{j=-N}^{N} \left[P_{12}^{\text{sep},\pm}(i, j; N) + P_{12}^{\text{sep},\pm}(j, i; N)\right] - P_{12}^{\text{sep},\pm}(i, i; N) \\ &= [P_{\downarrow}(i; N) + P_{\uparrow}(i; N)] - P_{12}^{\text{sep},\pm}(i, i; N). \end{aligned} \tag{18}$$

Given a one-particle probability distribution, say $P_{\downarrow}(i; N)$ [note that $P_{\uparrow}(i; N) = P_{\downarrow}(-i; N)$], our probability $\mathcal{P}^{\text{sep},\pm}(i; N)$ decreases with the joint probability $P_{12}^{\text{sep},\pm}(i, i; N)$. In fact, around the points $(40, 40)$ and $(-40, -40)$ in Fig. 2, we clearly have

$$\mathcal{P}^{-}(i; N) > \mathcal{P}^{\text{sep}}(i; N) > \mathcal{P}^{+}(i; N). \tag{19}$$

We see that, by introducing entanglement in the initial conditions of our two-particle quantum walk, the probability of finding at least one particle in a particular position on the line can actually be better or worse than in the case where the two particles are independent. Note that in this case, this does not depend on

the particular amount of entanglement introduced, as both states in Eq. (10) are maximally entangled, but rather on their symmetry/relative phase.

5. The Case of Mixed Initial State

Now, we turn to the discussion of whether a separable state can achieve the same results as entangled ones. First, we note that the differences between the cases of entangled and separable initial states occur with respect to the joint (two-particle) properties, rather than individual ones. We will show that a mixed separable state diagonal in the $\{|0,\downarrow\rangle_1|0,\uparrow\rangle_2, |0,\uparrow\rangle_1|0,\downarrow\rangle_2\}$ basis cannot achieve the same values for the correlation function and the distance between the two particles as entangled states. Let us consider the following mixed state, a weighted mixture of two alternatives described by the states $|0,\downarrow\rangle_1|0,\uparrow\rangle_2$ and $|0,\uparrow\rangle_1|0,\downarrow\rangle_2$, as a new initial state for our two-particle quantum walk:

$$\hat{\rho}_{12}^{MS}(0) = a|0,\downarrow\rangle\langle 0,\downarrow|_1 \otimes |0,\uparrow\rangle\langle 0,\uparrow|_2 + b|0,\uparrow\rangle\langle 0,\uparrow|_1 \otimes |0,\downarrow\rangle\langle 0,\downarrow|_2, \tag{20}$$

where $a, b \in \mathbb{R}^+$ such that $a + b = 1$. After N steps of the quantum walk, the system will be in the state $\hat{\rho}_{12}^{MS}(N) = \hat{U}^N \hat{\rho}_{12}^{MS}(0)\hat{U}^{\dagger N}$, and the marginal probability distributions will be

$$\begin{aligned} P_1^{MS}(i;N) &= aP_\downarrow(i;N) + bP_\uparrow(i;N), \\ P_2^{MS}(i;N) &= aP_\uparrow(i;N) + bP_\downarrow(i;N). \end{aligned} \tag{21}$$

We see that in general ($a \neq b$) they are biased, and that

$$\langle x_1\rangle = \sum_{i=-N}^{N} iP_1^{MS}(i;N) = a\langle x_\downarrow\rangle + b\langle x_\uparrow\rangle = (a-b)\langle x_\downarrow\rangle = -\langle x_2\rangle. \tag{22}$$

Here, we have used $\langle x_{\downarrow,\uparrow}\rangle \equiv \sum_{i=-N}^{N} iP_{\downarrow,\uparrow}(i;N)$, $P_{\downarrow,\uparrow}(i;N)$ being the probability distributions after N steps for initial states $|0,\downarrow\rangle$ and $|0,\uparrow\rangle$, and the fact that $\langle x_\downarrow\rangle = -\langle x_\uparrow\rangle$.

The joint probability distribution after N steps, as seen from Eq. (20), is given by

$$P_{12}^{MS}(i,j;N) = aP_\downarrow(i;N)P_\uparrow(j;N) + bP_\uparrow(i;N)P_\downarrow(j;N), \tag{23}$$

and is *not* the product of two one-particle marginal distributions $P_\downarrow(i;N)$ and $P_\uparrow(j;N)$. Our state $\hat{\rho}_{12}^{MS}(N)$ is separable, but it is also correlated. Its correlation function $C^{MS}(x_1,x_2)$ is not zero, as can clearly be seen from the definition of the correlation function (using $\langle x_1x_2\rangle = -\langle x_\downarrow\rangle^2$):

$$C^{MS}(x_1,x_2) \equiv \langle x_1x_2\rangle - \langle x_1\rangle\langle x_2\rangle = -\left[1-(a-b)^2\right]\langle x_\downarrow\rangle^2 = 2ab\langle x_\downarrow\rangle^2. \tag{24}$$

In Table 2, we present the values of $C^{MS}(x_1,x_2)$ for the case $a = b = \frac{1}{2}$. As the function $\left[1-(a-b)^2\right] \in [0,1]$, we see that $\hat{\rho}_{12}^{MS}(N)$ is always less correlated than a singlet entangled state. Actually, $\hat{\rho}_{12}^{MS}(N)$ is only a *classically correlated* state, and when compared with entangled states $|\psi_N^{\pm}\rangle_{12}$, it exhibits the same features as the pure separable state $|\psi_N^{PS}\rangle_{12}$. Namely, the expectation value of the distance is

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

the same for both pure and mixed separable states, $\langle \Delta_{12}^{MS} \rangle = \langle \Delta_{12}^{\text{sep}} \rangle$, as seen from the probability distribution (23). Similarly, we see that the probability of finding at least one particle in position i after N steps is the same for both pure separable and mixed separable case: $\mathcal{P}^{\text{sep}}(i;N) = \mathcal{P}^{MS}(i;N)$.

6. Conclusions, Further Study, and Implementations

There are a number of natural extensions to our work on discrete time quantum walks on a line with two entangled particles, starting with the study of other initial entangled states and the introduction of more particles. It would also be interesting to consider other graphs for the walks, as well as less standard coins, such as unbalanced or even entangling ones.[10,11] Another direction worth exploring would be a multiparticle quantum walk in continuous time, which does not use a coin, and which could be a relevant model to study the evolution of dilute quantum gases. Finally, it would also be of great interest to explore our ideas for particular applications or the design of efficient quantum algorithms. Random walks have been employed in such tasks as estimating the volume of a convex body[12] and the connectivity in P2P networks.[13] Quantum walks could have an important contribution to these problems, in particular our proposal with two entangled particles, which can cover more space than two independent particles, as we saw for the "−" case.

There have been several proposals to implement a single-particle quantum walk, using cavity QED,[14] optical lattices[15] and ion traps.[16] The latter could be adapted to the two-particle case by encoding the coin states in the electronic levels of two ions and the position in their *com* or *stretch* motional modes: the coin flipping could then be obtained with a $\pi/2$ Raman pulse and the shift with a conditional optical dipole force. Another possibility is to send two photons through a tree of balanced beam splitters which implement both the coin flipping and the conditional shift, again generalizing a scheme proposed for a single particle.[10,17] Note that this could be implemented with other particles as well, e.g. electrons, using a device equivalent to a beam splitter.[18] Finally, we would like to point out that, by considering indistinguishable bosons or fermions, the effects of quantum statistics can be used to prepare the initial entangled states, thus appearing again as a resource for quantum information processing.[19,20]

In this article, we introduced the concept of a quantum walk with two particles and studied it for the case of a discrete time walk on a line. Having more than one particle, we could now add a new feature to the walk: entanglement between particles. In particular, we considered initial states maximally entangled in the coin degrees of freedom and with opposite symmetries, and compared them to the case where the two particles were initially in a pure or mixed separable state (the particles being then independent or only classically correlated, respectively). We found that the entanglement in the coin states introduced spatial correlations between the particles, and that their average distance is larger in the "−" case than in the separable case, and is smaller in the "+" case. This could benefit

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

algorithmic applications which require two marked sites to be reached which are known *a priori* to be on the opposite or on the same sides of a line. We also found that the introduction of entanglement could increase or decrease the probability of finding at least one particle on a given point of the line. This increase could allow us to reach a marked site faster than with two unentangled quantum walkers. The entanglement in the initial conditions thus appears as a resource that we can tune according to our needs to enhance a given application (algorithmic or other) based on a quantum walk.

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

Acknowledgments

The authors would like to thank K. Banaszek, E. Kashefi and T. Schaetz for precious discussions. L. Sheridan gratefully acknowledges the support of the Nuffield Foundation through their Undergraduate Research Bursary. Y. Omar and N. Paunković gratefully acknowledge support from Fundação para a Ciência e a Tecnologia (Portugal), namely, through programs POCTI/POCI and CLC project POCI/MAT/55796/2004 QuantLog, partially funded by FEDER and the 3rd Community Support Framework of the European Social Fund. The research of Y. Omar and N. Paunković is also developed under the scope of the cooperation agreement 24P-22335/04 between CLE-UNICAMP and CLC-UTL. Y. Omar wishes to thank the Clarendon Laboratory for their hospitality. N. Paunković thanks the financial support of Elsag S.p.A. and of the European Commission, Contract No. IST-2001-39215 TOPQIP. Finally, Y. Omar and S. Bose would like to acknowledge support from the CRUP/BC collaboration grant B-22/05 and thank the Institute for Quantum Information at Caltech for their hospitality.

References

1. Y. Aharonov, L. Davidovich and N. Zagury, *Phys. Rev. A* **48** (1993) 1687.
2. J. Kempe, *Contemp. Phys.* **44** (2003) 307.
3. N. Shenvi, J. Kempe and K. B. Whaley, *Phys. Rev. A* **67** (2003) 052307.
4. A. Ambainis, quant-ph/0311001 (2003).
5. A. M. Childs *et al.*, in *Proc. 35th ACM Symp. Theory of Computing (STOC)*, 9–11 June, San Diego, CA (ACM, 2003), p. 59.
6. Y. Omar, N. Paunković, L. Sheridan and S. Bose, quant-ph/0411065 (2004).
7. D. Aharonov, A. Ambainis, J. Kempe and U. Vazirani, in *Proc. 33rd ACM Symp. on Theory of Computation (STOC)*, 22–26 June, Heraklion, Crete (ACM, 2001).
8. T. D. Mackay, S. D. Bartlett, L. T. Stephenson and B. C. Sanders, *J. Phys. A. Math. Gen.* **35** (2002) 2745.
9. B. Tregenna, W. Flanagan, R. Maile and V. Kendon, *New J. Phys.* **5** (2003) 83.
10. M. Hillery, J. Bergou and E. Feldman, *Phys. Rev. A* **68** (2003) 032314.
11. O. Buerschaper and K. Burnett, quant-ph/0406039 (2004).
12. M. Dyer, A. Frieze and R. Kannan, *J. ACM* **38** (1991) 1.
13. Q. Lv *et al.*, in *Proc. 16th ACM Int. Conf. Supercomputing (ICS)*, 22–26 June, New York City, NY (2002), pp. 84–95.
14. B. C. Sanders *et al.*, *Phys. Rev. A* **67** (2003) 042305.

15. W. Dür *et al.*, *Phys. Rev. A* **66** (2002) 052319.
16. B. C. Travaglione and G. J. Milburn, *Phys. Rev. A* **65** (2002) 032310.
17. H. Jeong, M. Paternostro and M. S. Kim, *Phys. Rev. A* **69** (2004) 012310.
18. R. C. Liu, B. Odom, Y. Yamamoto and S. Tarucha, *Nature* **391** (1998) 263.
19. Y. Omar, N. Paunković, S. Bose and V. Vedral, *Phys. Rev. A* **65** (2002) 062305.
20. N. Paunković, Y. Omar, S. Bose and V. Vedral, *Phys. Rev. Lett.* **88** (2002) 187903.

Int. J. Quantum Inform. 2006.04:573-583. Downloaded from www.worldscientific.com
by IMPERIAL COLLEGE LONDON on 04/21/15. For personal use only.

International Journal of Quantum Information
Vol. 4, No. 2 (2006) 365–370
© World Scientific Publishing Company

World Scientific
www.worldscientific.com

Int. J. Quantum Inform. 2006.04:365-370. Downloaded from www.worldscientific.com
by NANYANG TECHNOLOGICAL UNIVERSITY on 04/26/15. For personal use only.

ENTANGLEMENT-ASSISTED ORIENTATION IN SPACE

ČASLAV BRUKNER

Institut für Experimentalphysik, Universität Wien,
Boltzmanngasse 5, A-1090 Wien, Austria

Institut für Quantenoptik und Quanteninformation,
Österreichische Akademie der Wissenschaften,
Boltzmanngasse 3, A-1090 Wien, Austria

NIKOLA PAUNKOVIĆ

Fondazione I.S.I., Villa Gualino,
Viale Settimio Severo 65, 10133 Torino, Italy

TERRY RUDOLPH

Department of Physics, Blackett Laboratory,
Imperial College London, Prince Consort Road,
London SW7 2BW, UK

VLATKO VEDRAL

Institut für Experimentalphysik, Universität Wien,
Boltzmanngasse 5, A-1090 Wien, Austria

The School of Physics and Astronomy,
University of Leeds, Leeds, LS2 9JT, UK

Received 19 September 2005

We demonstrate that quantum entanglement can help separated individuals in making decisions if their goal is to find each other in the absence of any communication between them. We derive a Bell-like inequality that the efficiency of every classical solution for our problem has to obey, and demonstrate its violation by the quantum efficiency. This proves that no classical strategy can be more efficient than the quantum one.

Keywords: Quantum entanglement; communication complexity; spatial orientation.

Quantum entanglement is a phenomenon in which two or more quantum systems have to be described with reference to each other regardless of their spatial separation.[1–3,a] It leads to correlations that are inconsistent with local realism[6]

[a]Translation published in Refs. 4 and 5. A copy can be found at www.tu-harburg.de/rtz/rtz/it/QM/cat.html.

as demonstrated by violation of Bell's inequalities.[7] Although entanglement does not involve information transfer, it surprisingly can produce effects over arbitrary distances as if information had been transferred. It can *substitute or even eliminate any need of communication* that is classically necessary for achieving a goal of common interest of separated parties.[8–10] Entanglement can thus reduce the communication complexity of certain problems; the key ingredient for this is violation of local realism (quantum non-locality).[11]

Here, we show that entanglement can help individuals in making decisions if their goal is to find each other, even if there is no communication between them. This gives the problem properties of "pseudo-telepathy."[12] Reference 13 suggests a similar mechanism to explain cooperation of insects and demonstrates an advantage of the use of entanglement over some classical strategies. To provably demonstrate that no classical solution of our problem exists that can achieve the efficiency of the entanglement-based quantum solution, we derive a Bell-type inequality that every classical efficiency has to obey and demonstrate its violation by the quantum efficiency. In a broader context, we show that quantum non-locality — one of the most peculiar features of quantum physics — can be used to solve everyday if somewhat unusual problems, as similarly suggested by Refs. 14 and 15 and references therein.

We now define the problem in detail. We let two partners be on the two poles of the Earth, Alice on the North pole and Bob on the South pole. From either of the poles there are three possible paths, dark grey (1), light grey (2) and black (3), each with two possible directions, (+) and (−) (Fig. 1). Positive (negative) directions of each pair of the paths enclose an angle of 120° with each other if viewed from the poles. Each path is represented by one great circle on the globe.

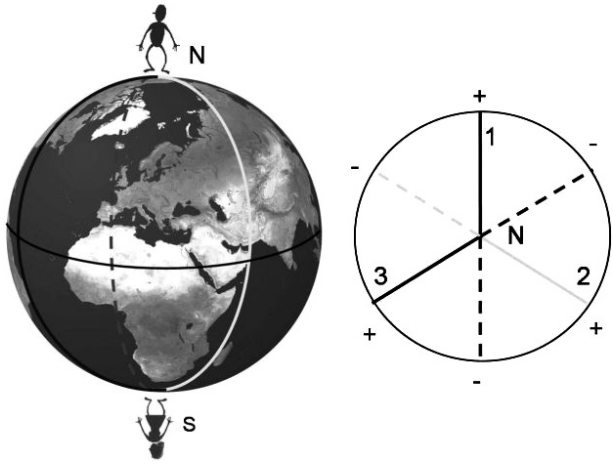


Fig. 1. Two partners are on the two poles of the Earth (left). From each pole there are three paths (dark grey 1, light grey 2 and black 3) and for each path there are two directions (+ and −) (right, view from the North pole). Which path and direction should the partners take to find each other at the equatorial line in the lack of any communication? (See text for details.)

Int. J. Quantum Inform. 2006.04:365-370. Downloaded from www.worldscientific.com
by NANYANG TECHNOLOGICAL UNIVERSITY on 04/26/15. For personal use only.

Int. J. Quantum Inform. 2006.04:365-370. Downloaded from www.worldscientific.com
by NANYANG TECHNOLOGICAL UNIVERSITY on 04/26/15. For personal use only.

With the aim to find each other, each partner chooses a path and direction and follows it along the great circle of the globe until she/he reaches the equatorial line. Arriving there, the partners certainly meet if they have taken the same path and the same direction (either $\{+,+\}$ or $\{-,-\}$). Alternatively, they may miss each other if they have taken either different paths or different directions of the same path. In the latter case, the partners will end at two opposite points of the equator. In the former case they will end at two points that are separated either under angle of 60°, in the case they have taken different directions (either $\{+,-\}$ or $\{-,+\}$), or under angle 120°, if they have taken the same directions (either $\{+,+\}$ or $\{-,-\}$). We will assume that both partners have a far-reaching view with an opening angle of 60° so that even if they took different paths but opposite directions they can see and subsequently find each other on the equator.

We assume that the partners cannot communicate nor that they have agreed in advance which path and direction to take in order to find each other. Furthermore, we assume that the choices of paths are completely independent, random and given with equal probabilities.[17,b] However, in choosing the directions Alice and Bob are allowed to use some previously shared classically correlated random strings or quantum entanglement. We therefore consider a situation in which Alice and Bob share no information whatsoever about which path each other's partner intends to take.[c]

In order to achieve their goal with the highest possible efficiency, Alice and Bob should maximize both the probability to take the same directions if they choose the same path and the probability to take opposite directions if they choose different paths. The overall probability of success is given by

$$P = \frac{1}{9}\left(\sum_{i=1}^{3} P_{ii}(\text{same}) + \sum_{i\neq j=1}^{3} P_{ij}(\text{opp}) \right), \tag{1}$$

where, e.g. $P_{ij}(\text{opp})$ is the probability that the partners take opposite directions if Alice chooses path i and Bob j. The factor $1/9$ is due to the assumption that every possible combination of the two paths taken by Alice and Bob is equally probable. We will now give the quantum solution of the problem. It is based on an example of quantum non-locality presented by Mermin.[16]

Suppose that the partners share a maximally entangled pair of photons in the state $|\phi^+\rangle = 1/\sqrt{2}(|H\rangle|H\rangle + |V\rangle|V\rangle)$, where $|H\rangle$ denotes horizontal polarization and $|V\rangle$ vertical polarization of a photon. Every partner chooses a path at random from the set $\{1,2,3\}$, independently of each other. The choice of paths determines a choice of directions of polarization measurements, given by the angles

[b]Later in the text we will see that this corresponds to the common assumption of experimenter's freedom in choosing between different possible measurement settings in the context of Bell's inequalities. See, for example, Refs. 17 and 18.

[c]Such a formulation of the problem invalidates a trivial possibility for partners to share a common set of instructions of the type "take black path, direction +."

$\{0°, 120°, -120°\}$: if the taken path is 1, the measurement angle is 0°, while in case of paths 2 and 3, measurement angles are 120° and −120°, respectively. Polarization measurements are performed on each partner's photons. If the outcome is H, the partner takes direction +, and if it is V she/he takes direction −. In 3/9 fraction of cases, the partners choose the same measurement directions and obtain the same outcomes with certainty. In the remaining 6/9 fraction of cases, they choose different measurement directions, in which case the probability of them obtaining opposite results is 3/4. Thus, one has $P = 5/6 \approx 83\%$ for the success rate in the quantum protocol.

In a classical protocol, instead of an entangled pair of photons, Alice and Bob can use some previously distributed classically correlated random strings of variables which may improve the success of the protocol. On the basis of these strings, they could make their decision which directions (+ or −) to choose, but the choice of the path (dark grey, light grey, or black) is, as already mentioned, assumed not to be dependent on the strings.

The best classical strategy is successful in $7/9 \approx 78\%$ of the cases. This can be achieved in a model where Alice and Bob share classically correlated variables and in which each variable carries an internal code, determining whether the outcome H or V will emerge, for each of the three possible choices of paths 1, 2 and 3. Thus, the set of all possible codes for both variables given to Alice and Bob is: $\{(HHH), (VVV), (HHV), (HVH), (VHH), (HVV), (VHV), (VVH)\}$. Here the position in each code plays the role of the measurement angle, while the corresponding value (H or V) at that position corresponds to a measurement outcome. Imagine that each pair of variables, one given to Alice and the other to Bob, carries the same code, so as to have the maximal possible probability of success in the case when both partners choose the same path: $\frac{1}{9}\sum_{i=1}^{3} P_{ii}(\text{same}) = 3/9$. For any of those pairs of codes, if the choices of paths are different, then the probability of emerging opposite outcomes is at most 2/3, which can be easily checked by simple inspection (the value 7/9 is obtained if neither of codes $\{(HHH), (VVV)\}$ is used). Thus, the efficiency is $P = (3/9) + (6/9)(2/3) = 7/9$ in this protocol.

No classical protocol can be more efficient than this, in the lack of any communication between the two parties. This is based on the proof given below that the combination of probabilities as given in Eq. (1) satisfies a Bell inequality with 7/9 being a local realistic bound. Thus, the probability of success of any classical protocol is bounded, while those of quantum protocols can exceed the limit by utilizing quantum non-locality.

We now give the proof that

$$\sum_{i=1}^{3} P_{ii}(\text{same}) + \sum_{i\neq j=1}^{3} P_{ij}(\text{opp}) \leq 7, \tag{2}$$

for all local realistic models. This new Bell's inequality with three possible measurement settings per observer is a byproduct of our analysis. Consider two observers, Alice and Bob, and allow each of them to choose between three dichotomic

Int. J. Quantum Inform. 2006.04:365-370. Downloaded from www.worldscientific.com
by NANYANG TECHNOLOGICAL UNIVERSITY on 04/26/15. For personal use only.

Int. J. Quantum Inform. 2006.04:365-370. Downloaded from www.worldscientific.com
by NANYANG TECHNOLOGICAL UNIVERSITY on 04/26/15. For personal use only.

observables, determined by some local parameters denoted here a_1, a_2 and a_3 for Alice and b_1, b_2 and b_3 for Bob. The assumption of local realism implies the existence of three numbers A_1, A_2 and A_3 for Alice and B_1, B_2 and B_3 for Bob, where the numbers take values $+1$ or -1 and describe the predetermined results of corresponding measurements. In other words, for given specific numerical values, $(A_1A_2A_3)$ is a code from our previous example, while the values $+1$ and -1 play the role of H and V, respectively. In a specific run of the experiment, the correlations between two observations can be represented by the product of the type A_iB_j. The correlation function is then the average over many runs of the experiment $E(a_i, b_j) = \langle A_iB_j\rangle$.

The following combination of the predetermined results has a maximal value as given by

$$\max\{A_1(B_1 - B_2 - B_3) + A_2(B_2 - B_1 - B_3) + A_3(B_3 - B_1 - B_2)\} = 5. \quad (3)$$

After averaging this expression over the ensemble of the runs of the experiment, one obtains the following Bell inequality:

$$\left| \sum_{i=1}^{3} E(a_i, b_i) - \sum_{i\neq j}^{3} E(a_i, b_j) \right| \leq 5. \quad (4)$$

Using the connection between correlation functions and probabilities[d], $E(a_i, b_i) = 2P_{ii}(\text{same}) - 1$ and $E(a_i, b_j) = 1 - 2P_{ij}(\text{opp})$ when $i \neq j$, one finally obtains inequality (2). Its local realistic limit demonstrates that the efficiency of the particular classical solution discussed above is the optimal one. The quantum solution is based on violation of the inequality by a factor 7.5 with the quantum entangled state $|\phi^+\rangle$.

Our problem can be seen as an instance studied in the field of communication complexity. Typically, in communication complexity problems, two spatially separated parties receive local input data, e.g. one party receives number i and the other j. Their goal is to compute a given function $f(i, j)$. In one class of these problems, only a restricted amount of communication between the parties is allowed, so that in general they cannot arrive at the correct value of the function with certainty. While an error is allowed, the parties try to compute the function correctly with as high probability as possible.

In our example, no communication between Alice and Bob is allowed. The input data i and j correspond to Alice's and Bob's choices of paths, respectively. Function $f(i, j)$ is defined in the following way: $f(i, i) = 1$ and $f(i, j) = -1$ if $i \neq j$. On the basis of i and j Alice and Bob produce new local data A and B (obtained as results $+1$ or -1 of local measurements of polarization along directions i and j in quantum protocol) that define the directions of their movements along paths

[d] $E(a_i, b_j) = \langle A_iB_j\rangle = \sum_{A_i,B_j\in\{-1,1\}} P(A_i, B_j)A_iB_j = P_{ij}(\text{same}) - P_{ij}(\text{opp}) = 2P_{ij}(\text{same}) - 1 = 1 - 2P_{ij}(\text{opp})$, as $P_{ij}(\text{same}) + P_{ij}(\text{opp}) = 1$. Here $P(A_i, B_j)$ is probability that Alice obtains A_i and Bob B_j, for the choices of paths i and j, respectively.

i and j, respectively. Only if $f(x, y) = A \cdot B$ will they find each other (in which case they also compute the function correctly).

In conclusion, we have demonstrated that quantum entanglement can help separated individuals to orient themselves on a sphere if their aim is to find each other, even in the absence of any communication between them. In future, it will be interesting to investigate whether there are further problems in life science, economics or every-day situations whose solutions favor quantum correlations over classical ones.

Int. J. Quantum Inform. 2006.04:365-370. Downloaded from www.worldscientific.com
by NANYANG TECHNOLOGICAL UNIVERSITY on 04/26/15. For personal use only.

Acknowledgments

Časlav Brukner has been supported by the Austrian Science Foundation (FWF) Project SFB 1506 and the European Commission, the Integrated Project Qubit Applications (QAP), Contract Number 015848. Nikola Paunković is funded by the European Commission, Contract No. IST-2001-39215 TOPQIP. Terry Rudolph is supported by the Engineering and Physical Sciences Research Council.

The authors gratefully acknowledge discussions with Karl Svozil and Johann Summhammer and hospitality of Oxford's "The Three Goats' Heads" pub and Vienna coffee house "Prückel."

References

1. E. Schrödinger, *Naturwissenschaften* **23** (1935) 807–812.
2. E. Schrödinger, *Naturwissenschaften* **23** (1935) 823–828.
3. E. Schrödinger, *Naturwissenschaften* **23** (1935) 844–849.
4. E. Schrödinger, *Proc. Am. Phil. Soc.* **124** (1935) 323–338.
5. E. Schrödinger, in *Quantum Theory and Measurement*, eds. J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton), pp. 152–167.
6. A. Einstein, B. Podolsky and N. Rosen, *Phys. Rev.* **47** (1935) 777.
7. J. Bell, *Physics (Long Island City, N.Y.)* **1** (1964) 195.
8. A. C.-C. Yao, in *Proc. 11th Annual ACM Symp. Theory of Computing* (1979), pp. 209–213.
9. R. Cleve and H. Buhrman, *Phys. Rev. A* **56** (1997) 1201.
10. G. Brassard, quant-ph/0101005 (2001).
11. Č. Brukner, M. Zukowski, J.-W. Pan and A. Zeilinger, *Phys. Rev. Lett.* **92** (2004) 127901.
12. G. Brassard, A. Broadbent and A. Tapp, in *Proc. 8th Int. Workshop on Algorithms and Data Structures*, Lecture Notes in Computer Science, Vol. 2748 (Springer-Verlag, 2003), pp. 1–11.
13. J. Summhammer, quant-ph/0503136 (2005).
14. A. M. Steane and W. van Dam, *Phys. Today* **53** (2000) 35–39.
15. K. Jacobs and H. Wiseman, *Am. J. Phys.* **73** (2005) 932–937.
16. N. D. Mermin, *Am. J. Phys.* **49** (1981) 940.
17. J. S. Bell, Free variables and local causality, *Dialectica* **39** (1985) 103–106.
18. J. Kofler, T. Paterek and Č. Brukner, *Phys. Rev. A* **73** (2006) 022104.

A Network Server for Distributing Quantum Random Numbers

Nuno A. Silva[*], **Maurício J. Ferreira**[*†], **André Carvalho**[*], **André Souto**[‡§], **Nikola Paunković**[‡], **Paulo Mateus**[‡¶], **António Teixeira**[*†], **Armando N. Pinto**[*†]

[*]*Instituto de Telecomunicações, University of Aveiro, 3810-193, Aveiro, Portugal*

[†]*Department of Electronics, Telecommunications and Informatics, University of Aveiro, 3810-193, Aveiro, Portugal*

[‡]*Instituto de Telecomunicações, 1049-001 Lisbon, Portugal*

[§]*LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal*

[¶]*Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

email: nasilva@ua.pt

2023 23rd International Conference on Transparent Optical Networks (ICTON) | 979-8-3503-0303-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICTON59386.2023.10207372

ABSTRACT

Quantum random number generators (QRNGs) based on homodyne detection have a potential to be a cost-effective solution to deploy random numbers at high-speed rates over networks. The new information technologies, such as quantum key distribution, cloud computing, and big data, demands true randomness for those cryptosystems to be secure and private. This is incompatible with known security loopholes of classical pseudo-random number generators. In this work, we discuss our recent advances in the implementation of a vacuum-based QRNG service to be used in a distributed network environment. We will also highlight the advantages of using quantum technologies to generate random numbers in comparison with classical approaches.

Keywords: Random Number Generation, Vacuum Fluctuations, Multiplexing scheme, Min-entropy Optimization

1. Introduction

Random Numbers (RNs) are essential for many new information technologies such as quantum key distribution, cloud computing, and big data [1]. They are particularly indispensable for cryptographic systems since, to remain secure and private, these demand true randomness [2], incompatible with the deterministic nature of traditional RN Generators (RNGs). These Pseudo RNGs (PRNGs) have been widely adopted due to their easiness of implementation and high throughput rates. Unfortunately, they are predictable to an adversary with enough computational resources [3], which is especially critical with the maturation of quantum computing. Quantum RNGs (QRNGs) harness the probabilistic nature of quantum phenomena to obtain true RNs and thus address the security loopholes of PRNGs [1]. Promising randomness sources have been proposed. Among them, cost-effective QRNGs based on homodyne detection of vacuum fluctuations allow the generation of RNs at high-speed rates [4–6]. Recent efforts focused mainly on achieving chip-integration [7], modeling additional experimental imperfections [5], or semi-device independent implementation [8]. Additionally, an increase of the effective real-time generation rate was attempted by selecting independent modes of the vacuum state to generate subentropy sources and thus remove randomness extraction bottlenecks [9]. Unfortunately, the experimental limitations of such a method were not studied.

In this article, we analyze the impact of practical limitations on an entropy optimization technique for vacuum-based QRNGs that relies on multiplexing and amplifying independent side-band frequency modes of the vacuum state to obtain a set of parallel RN generation channels. We simulate the required electronic components and assess its viability for a previously demonstrated experimental scheme and obtain figures such as the expected channel entropy.Finally, we also discuss the implementation of a QRNG service into a distributed network environment.

2. A Multiplexed QRNG based on Vacuum Fluctuations

A schematic representation of the proposed multiplexed vacuum-based QRNG is presented in Fig 1. Here, only the Balanced Homodyne Detection (BHD) stage was experimentally realized, while the performance of the multiplexing scheme was assessed through a numerical simulation. The experimental implementation consists of a 1550.92 nm continuous-wave laser at 11 dBm acting as a Local Oscillator (LO). It interacts with the vacuum state at the input port of a balanced Beam Splitter (BS), and the output beams are detected by a balanced receiver (WL-BPD1GA) with an output bandwidth of 1 GHz. Here, the resulting photocurrents are subtracted and amplified by a Transimpedance Amplifier (TIA) such that a signal proportional to the normally-distributed amplitude quadrature of the vacuum state is obtained [4]. Subsequently, the detector's response is, at this stage, sampled at 983.04 MSa/s by a 16-bit Analog-to-digital Converter (ADC) module (Texas Instruments ADS54J60EVM) with a ±0.95 V acquisition range.

Authorized licensed use limited to: b-on: Universidade de Lisboa Reitoria. Downloaded on September 19,2025 at 17:20:52 UTC from IEEE Xplore. Restrictions apply.

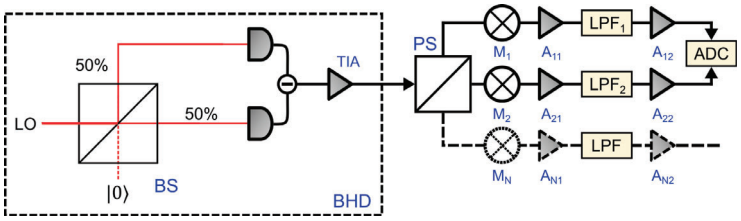


Figure 1: Proposed multiplexed QRNG: the BHD output signal is split by a PS into multiple channels. Each frequency sidemode is then selected by a frequency mixer (M_N) and a LPF. Finally, each channel is amplified to optimize the entropy available.

The throughput of the scheme is limited by its sampling rate and an entropy estimation that quantifies the impact of side-information introduced by non-random noise [10]. A common approach is to consider the worst-case conditional min-entropy, where it is possible to find an optimized ADC acquisition range that minimizes the number of underpopulated or over-saturated bins and yields an optimal min-entropy estimation [11]. In this scheme, the high Quantum-to-classical Noise Ratio (QCNR) of 11.7 dB observed yields the extraction of 8.39 random bits from each measurement, which supports output rates up to 8.23 Gbps. Nonetheless, an optimized min-entropy of 13.97 bits per sample is theoretically achievable by adjusting the ADC acquisition range. Alternatively, the output signal could be amplified by approximately 31.58 dB to match this optimal range level. Unfortunately, the finite Gain-Bandwidth Product (GBP) of the amplifiers limits the scheme's generation rate since, given by the Nyquist theorem, the sampling rate chosen must be less than twice the signal's bandwidth to avoid randomness overlapping [4].

The multiplexing stage selects non-overlapping sideband frequency modes of the quantum vacuum state within the detector's bandwidth. These are independent and thus constitute subentropy sources that generate parallel random generation channels that can be sampled at lower rates [9]. In practice, it consists of a Power Splitter (PS), which divides the output signal into an arbitrary number of channels. Each subsequent output is then mixed down in a frequency mixer (M_N) with an rf signal whose frequency f_s defines the center of the selected spectral sideband, and is filtered through a Low-pass Filter (LPF) with a certain cut-off frequency f_c. Finally, each channel is amplified to the corresponding optimization level by two amplifiers (A_{N1} and A_{N2}) of equal gain, respectively applied before and after the spectral selection. Here, the case of a dual channel was considered. The first was mixed down with a 200 MHz sinusoidal wave, while the second component was used to sweep central frequencies up to 600 MHz. Moreover, an almost-ideal 40th order LPF with f_c of 100 MHz was initially implemented, and a low-noise operational amplifier with noise power density of 0.98 nV/$\sqrt{Hz}$ was selected to implement the two-stage amplifier.

3. Results and Discussion

In Fig. 2(a) the conditional min-entropy is presented as a function of the gain introduced by each amplifier stage for a GBP of 1 GHz. A gain of approximately 28 dB is required to reach entropy optimization. Meanwhile, only 26 dB would be necessary if a GBP of 8 GHz were considered. This is explained by the bandwidth reduction imposed

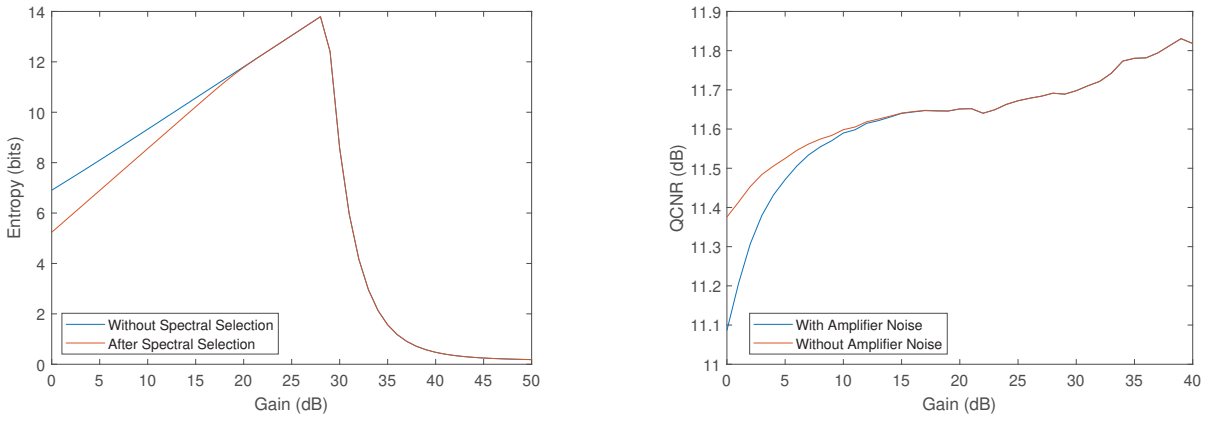
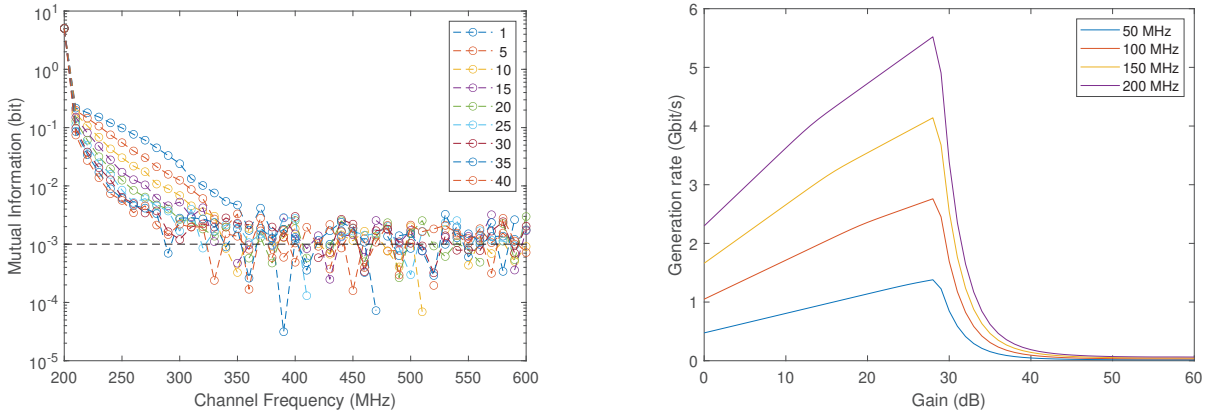


Figure 2: a) Worst-case conditional min-entropy as a function of the gain of each amplifier for a GBP of 1 GHz. b) QCNR as a function of the amplifier gain (without) considering the A_{N2} excess noise.

Authorized licensed use limited to: b-on: Universidade de Lisboa Reitoria. Downloaded on September 19,2025 at 17:20:52 UTC from IEEE Xplore. Restrictions apply.

for larger channel gains and can be seen when the spectral selection stage is removed. In this case, the effective bandwidth of the output signal is no longer limited to the 100 MHz from the LPF and its observed variance thus increases. Consequently, lower gain factors are required to reach the same entropy values. This difference disappears for gains larger than 20 dB since this is the level at which the bandwidth imposed by the amplifiers equals f_c and the two cases become equivalent. Although this could be minimized by additional amplification stages, the wide ADC acquisition range thus imposes the use of high GBPs. Moreover, the expected QCNR decrease due to additional noise from the second amplifier, A_{N2}, is depicted in Fig. 2(b). As can be seen, its impact is limited for small gain factors due to the low noise amplifier considered, and a maximum reduction of 0.6 dB is expected. As a consequence of the two-stage scheme explored, a small decrease of 0.3 dB is still seen even when the excess noise of the second stage is not considered. Nonetheless, by increasing the channel amplification, the reduction of the effective bandwidth allows this figure to increase beyond the value calculated using the entire bandwidth of the detector.

With two sequences of 1 M samples, the mutual information was also calculated for increasing channel separation and LPF orders. As seen in Fig. 3(a), for close-to-ideal filters, this quantity rapidly converges to the value expected for independent sequences of this size. This validates the proposed scheme as an adequate method to obtain parallel subentropy sources within the homodyne detector's bandwidth. Despite this, for realistic filter order values, side-information can still be present up to twice the channel bandwidth, which should be accounted for when designing a practical implementation. Finally, the expected generation rate for each channel was assessed as a function of f_c. In Fig. 3(b), this figure is calculated for each of the gain factors considered. As represented, smaller channel bandwidths require higher channel amplification levels to reach an optimized min-entropy. Nevertheless, although an average of 13.79 bits was obtained across all f_c values, the throughput increases proportionally with the maximum sampling rate. For the 8 GHz GBP selected, this limit is twice f_c for all amplification factors considered.



(a) (b)

Figure 3: a) Mutual information between two sequences of 1 M samples as a function of the f_s from the second channel. In each case, increasing LPF orders were evaluated. b) Expected channel generation rate for different LPF cut-off frequencies as a function of the amplification level. Here, a GBP of 8 GHz was selected.

According to the previous analysis, the dual-channel configuration explored only allows generation rates up to 5.56 Gbps, and thus no performance gain is actually obtained. Nonetheless, this multiplexing technique relaxes the performance requirements for the physical devices by lowering the required sampling rate, which is important to reach low-cost high-speed implementations. Moreover, the chosen f_c supports up to 4 channels within the detector's bandwidth. Assuming that the available entropy remains constant between channels, an additional third or fourth channel would increase this value to 8.33 Gbps and 11.11 Gbps, respectively. This corresponds to increases of 1.21% and 34.99% in comparison with the experimentally implemented scheme. Nonetheless, since the frequency sidebands must be properly spaced, some f_c values cannot support the maximum number of channels allowed. In these cases, larger bandwidths will lead to a larger decrease in the overall throughput. It thus appears that, without careful consideration of the values chosen, lower bandwidths allow for better maximization of the overall generation rates.

4. A Network Server for Distributing Quantum Random Numbers

A schematic representation of the proposed network server to distribute RNs through internet connection is present in Fig. 4. In it we observe that the proposed network server is divided in two main blocks, i.e. a back-end and a

Authorized licensed use limited to: b-on: Universidade de Lisboa Reitoria. Downloaded on September 19,2025 at 17:20:52 UTC from IEEE Xplore. Restrictions apply.

front-end. The latter is the web page that users see, whereas the back-end is responsible to provide the front-end with all the data including authorization and authentication. It is here that the physical QRNG is connected in a sensor pol where several different RNGs could be connected. To deal with the possibility of high demand from the front-end for short to medium periods the sensor pol is connected to the service where a cache/buffer memory where sequences of RNG can be stored. To guarantee the necessary authorization for the front-end users we also develop a user service that manages the information about the users and their respective authentication. Besides that in our back-end we also develop an API gateway that connects developers to the services through a unique API. Finally, the back-end also contains a discover service that provides load balancing capabilities and improves design by removing direct endpoints and translating that to a name which will be converted to an endpoint inside this service. Connected to all of those services is the front-end where the users can have access to the RNs on demand.

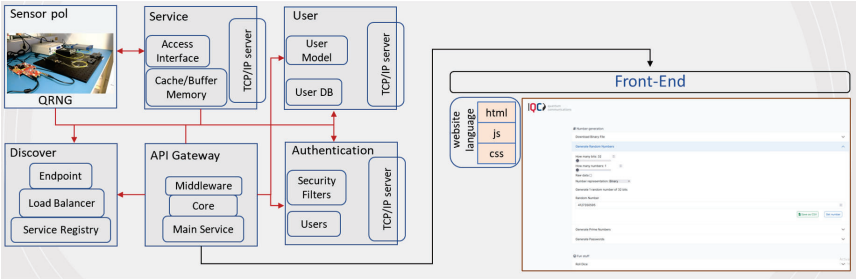


Figure 4: Schematic diagram of the network server.

5. Conclusions

In conclusion, the viability of an entropy optimization technique that derives multiple randomness generation channels from a traditional vacuum-based QRNG was assessed for the output of an experimental implementation in a dual-channel configuration. By numerically simulating the required electronic components, the expected conditional quantum min-entropy of each channel was calculated for increasing amplification levels. For the 100 MHz channel bandwidth chosen, results show an expected 64.36 % increase, although the imposed reduction in the sampling rate decreases the generator's throughput to 5.56 Gbps. Nonetheless, additional generation channels are supported by the chosen bandwidth, which would allow us to obtain increases in the generation rate up to 34.99%. Finally, we briefly introduced the server structure necessary to implement the QRNG service in a distributed network environment.

ACKNOWLEDGEMENTS

This work was supported in part by Fundação para a Ciência e a Tecnologia (FCT) through national funds, by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme (COMPETE 2020) of the Portugal 2020 framework, under the project UIDB/50008/2020-UIDP/50008/2020 and QuantumPrime (PTDC/EEI-TEL/8017/2020). M. F. acknowledges the 2022.09584.BD PhD grant from FCT. A.S. acknowledges funds granted to LASIGE Research Unit, ref. UID/CEC/00408/2013. N.P. acknowledges FCT project CERN/FIS-PAR/0023/2019, and FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT000.

REFERENCES

[1] M. Herrero-Collantes and J. C. Garcia-Escartin, "Quantum random number generators," *Rev. Mod. Phys.*, vol. 89, p. 015004, Feb 2017.
[2] J. Bouda, M. Pivoluska, M. Plesch, and C. Wilmott, "Weak randomness seriously limits the security of quantum key distribution," *Phys. Rev. A*, vol. 86, p. 062308, Dec 2012.
[3] J. Kelsey, B. Schneier, D. Wagner, and C. Hall, "Cryptanalytic attacks on pseudorandom number generators," in *Fast Software Encryption*, S. Vaudenay, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 168–188.
[4] M. J. Ferreira, N. A. Silva, A. N. Pinto, and N. J. Muga, "Characterization of a Quantum Random Number Generator Based on Vacuum Fluctuations," *Applied Sciences*, vol. 11, no. 16, 2021.
[5] C. Bruynsteen, T. Gehring, C. Lupo, J. Bauwelinck, and X. Yin, "100-Gbit/s Integrated Quantum Random Number Generator Based on Vacuum Fluctuations," *PRX Quantum*, vol. 4, p. 010330, Mar 2023.
[6] T. Gehring, C. Lupo, A. Kordts, D. Solar Nikolic, N. Jain, T. Rydberg, T. Pedersen, S. Pirandola, and U. Andersen, "Homodyne-based quantum random number generator at 2.9 Gbps secure against quantum side-information," *Nature Communications*, vol. 12, 01 2021.
[7] B. Bai, J. Huang, G.-R. Qiao, Y.-Q. Nie, W. Tang, T. Chu, J. Zhang, and J.-W. Pan, "18.8 Gbps real-time quantum random number generator with a photonic integrated chip," *Applied Physics Letters*, vol. 118, no. 26, p. 264001, 2021.
[8] M. Avesani, D. G. Marangon, G. Vallone, and P. Villoresi, "Source-device-independent heterodyne-based quantum random number generator at 17 Gbps," *Nature Communications*, vol. 9, no. 1, pp. 1–7, 2018.
[9] X. Guo, C. Cheng, M. Wu, Q. Gao, P. Li, and Y. Guo, "Parallel real-time quantum random number generator," *Opt. Lett.*, vol. 44, no. 22, pp. 5566–5569, Nov 2019.
[10] M. Ferreira, N. Silva, and N. Muga, "Efficient Randomness Extraction in Quantum Random Number Generators," in *Anais do II Workshop de Comunicação e Computação Quântica*. Porto Alegre, RS, Brasil: SBC, 2022, pp. 31–36.
[11] J. Y. Haw, S. M. Assad, A. M. Lance, N. H. Y. Ng, V. Sharma, P. K. Lam, and T. Symul, "Maximization of extractable randomness in a quantum random-number generator," *Phys. Rev. Applied*, vol. 3, p. 054004, May 2015.

Authorized licensed use limited to: b-on: Universidade de Lisboa Reitoria. Downloaded on September 19,2025 at 17:20:52 UTC from IEEE Xplore. Restrictions apply.

On minimal assumptions to obtain a universally composable quantum bit commitment

M. Lemus[1,2], **P. Yadav**[1,2], **P. Mateus**[1,2], **N.Paunković**[1,2], **A. Souto**[3,4,2]
[1]*Instituto Superior Técnico ULisboa, Av. Rovisco Pais, Lisboa, Portugal*
[2]*Instituto de Telecomunicações, Av. Rovisco Pais, Lisboa, Portugal*
[3]*Faculdade Ciências ULisboa, Campo Grande, Lisboa, Portugal*
[4]*LASIGE, Dep. de Informática Faculdade Ciências ULisboa, Campo Grande, Lisboa, Portugal*
e-mail: {mariano.lemus, preeti.yadav, pmateus, npaunkovic}@tecnico.ulisboa.pt, ansouto@fc.ul.pt

ABSTRACT
It is well known, due to the no-go theorems, that perfectly secure bit commitment is impossible. For that reason, there exist many proposals for bit commitment protocols that are secure under certain assumptions (such as the existence of a trusted party, computational hardness, or noisy memories). In this paper, we investigate the minimal assumptions required to obtain a secure bit commitment protocol. We show that it is possible to construct a bit commitment scheme with two parties Alice and Bob as long as there is a third party that is willing to broadcast either entangled qubits for the parties or to produce classical strings that are xor-correlated. We also show that the bit commitment scheme proposed is universally composable.
Keywords: Bit commitment schemes, Universal compasability.

1. INTRODUCTION

The idea of a commitment scheme was first proposed by Blum [1] in 1983, in a protocol where two parties can play fairly a coin tossing game over the phone based on the existence of one-way functions. Since then, enough work has been done that manifests its importance as a primitive for several important cryptographic tasks namely zero-knowledge proofs, verifiable secret-sharing, electronic-voting, digital signatures, among others [2]. Furthermore, in the context of quantum cryptography, bit commitment allows for the realization of oblivious transfer [3], which in turn enables general secure multiparty computation [4].

Unfortunately, in the framework of classical cryptography, it was shown that commitment schemes cannot be securely realized without assuming the use of additional resources [5]. Moreover, the impossibility of designing bit commitment and oblivious transfer schemes that are unconditionally secure in the quantum realm was proven [6], [7]. In particular, no quantum bit commitment protocol can be both concealing and binding unconditionally. Due to this significant drawback, the scientific community started to consider restrictions on the power of the adversaries to design practical commitment schemes.

In this paper, we consider the following question: *"What are the minimal assumptions that one can have to design a perfectly secure bit commitment that is universally composable?"* Notice that since the bit commitment is usually used as a subroutine of more complex protocols, the universal composability (i.e., security in any computational environment) is mandatory, in the sense that only with this property such a scheme is useful. In this work, we show that one can achieve a perfectly secure bit commitment by using a third party beamer that distributes asymmetric information between the parties involved. We say that this assumption is minimal because all the associated algorithm time complexities, as well as the communication complexities, are linear in the security parameter, and the beamer is non-interactive.

This work is organized as follows. In Section 2, we formally define the concept of bit/string commitment. In Section 3, we characterize the Assymetric Quantum Beamer resource, to be used for our bit commitment protocol, described in Section 4, whose composable security is then analyzed. Finally, we state the conclusions of this work in Section 5.

2. PRELIMINARIES

A commitment scheme involves two parties: the sender Alice, and the receiver Bob. It consists of two stages, the *commitment phase*, and the *opening phase*. The goal is for Alice to commit to a certain secret value during the commitment phase, which is revealed to Bob in the opening phase. Figure 1 shows the description of the commitment functionality $\mathcal{F}_{\text{COM}}$. Intuitively, a commitment protocol is called secure if it holds the following properties:

1) *Binding:* Alice cannot open to a different value than the one she initially committed to.
2) *Concealing:* Bob cannot get information about the committed value before the opening phase.

When the size k of the committed value is one bit, we call it a bit commitment (BC) functionality $\mathcal{F}_{\text{BC}}$. Having access to secure BC is enough to securely realize commitments of strings of arbitrary size. Therefore, in this work, we focus only on the problem of securely realizing BC.

The security of cryptographic protocols can be studied from two perspectives, stand-alone security and composable security. A stand-alone security proof consists in showing that the main properties of the intended

978-1-7281-2779-8/19/$31.00 ©2019 IEEE

Authorized licensed use limited to: University of Exeter. Downloaded on June 14,2020 at 15:21:39 UTC from IEEE Xplore. Restrictions apply.

Functionality $\mathcal{F}_{\text{COM}}$

Parameters:

- Parties Alice and Bob.
- Size k of the committed value.

1) Upon receiving an input (*commit, x*) from Alice, if no value has previously been committed, output the message (*committed*) to Bob.
2) Upon receiving the input (*open*) from Alice, if a value x has previously been committed, output the message (*open, x*) to Bob, and halt.

Figure 1: Commitment functionality.

service hold. In the case of BC, it is required that the soundness (the outputs are correct whenever the protocol finishes without errors), binding and concealing properties hold. On the other hand, a composable security proof consists in showing that the parties running the protocol, while taken together as a single bigger party, are indistinguishable from a simulated machine running the ideal functionality [8]. When they are indistinguishable only in computationally bounded environments, we say that the protocol is computationally secure. If they are indistinguishable in any environment, we say that the protocol is perfectly secure. In general, protocols are secure up to a negligible probability given by a security parameter. In these cases, perfect security is also referred to as statistical security.

It has been shown that secure BC is impossible without assuming access to additional resources. Notably, channel authentication between the parties is a necessary, but not sufficient resource. Computationally secure BC protocols have been demonstrated in the Common Reference String, Random Oracle, and Public Key Infrastructure models [9]. A statistically secure BC scheme in the Binary Symmetric Channel model has been proposed [10]. In this work, we aim to describe a non-interactive resource, from which a composable and statistically secure BC protocol can be realized.

3. ASYMMETRIC QUANTUM BEAMER MODEL

In the Asymmetric Quantum Beamer (AQB) model, we assume that each pair of participants have access to a non-interactive service which constantly outputs messages to them without inputs from any party. Figure 2 shows a description of the $\mathcal{F}_{\text{AQB}}$ functionality, which sends classical strings to one party and a combination of classical strings and a quantum state to the other. We assume that the communication channels between $\mathcal{F}_{\text{AQB}}$ and the parties are private and authenticated. Furthermore the function f and the basis set $\mathcal{B}$ should be chosen such that, given s, computing $f(s)$ and $f(\tilde{s})$ can be done efficiently, as well constructing the basis $\mathcal{B}_s$.

Functionality $\mathcal{F}_{\text{AQB}}$

Parameters:

- Parties Alice and Bob.
- Security parameter m.
- Injective function $\mathcal{C} : \{0,1\}^{m/2} \to \{0,1\}^m$, such that $\mathcal{C}(s) = s.\, f(s)$, where f is an (almost) perfectly non-linear function.
- Time interval Δt.
- Set $\mathcal{B} = \{(|e_0^s\rangle, |e_1^s\rangle)\}_{s\in\{0,1\}^{m/2}}$ of bases of $\mathbb{C}^2$, such that $\forall s\ \exists \tilde{s}\ (|e_0^s\rangle, |e_1^s\rangle) = (|e_1^{\tilde{s}}\rangle, |e_0^{\tilde{s}}\rangle)$.

Upon activation, $\mathcal{F}_{\text{AQB}}$ sets the time index $\tau = 1$ and repeats the following steps at regular time intervals Δt.

1) Sample random $s \in \{0,1\}^{m/2}$, $\ell \in \{0,1\}^m$, and bit j.
2) Compute $\mathcal{C}(s)$.
3) Output $(\mathcal{C}(s) \oplus \ell, j, \tau)$ to Alice and $(\ell, |e_j^s\rangle, \tau)$ to Bob.
4) Increase τ by 1.

Figure 2: Assymetric quantum beamer functionality.

4. PROTOCOL

In this section, we describe a protocol in the AQB model, requiring only classical (authenticated) communication between the parties and local quantum measurement capabilities for one party. In the bit commitment protocol $\pi_{\text{BC}}^{\text{AQB}}$, shown in Fig. 3, the aim of Alice is to send to Bob a bit value b she wants to commit to. Due to the concealing requirement, it has to be encrypted, in our case by the key bit j given to Alice by $\mathcal{F}_{\text{AQB}}$, i.e., she sends $b \oplus j$. In order for Bob to recover the value b, he needs to obtain the encryption key j, by measuring the state $|e_j^s\rangle$ (given to him by $\mathcal{F}_{\text{AQB}}$) in the appropriate basis $\mathcal{B}_s$. During the opening phase, Alice provides Bob

Authorized licensed use limited to: University of Exeter. Downloaded on June 14,2020 at 15:21:39 UTC from IEEE Xplore. Restrictions apply.

with sufficient information to do so. She sends him the encrypted value of the basis label, given by $\mathcal{C}(s) \oplus \ell$, which he can decrypt, having its encryption key ℓ, the former being given by $\mathcal{F}_{\text{AQB}}$ to Alice, and the latter to Bob.

Protocol $\pi_{\text{BC}}^{\text{AQB}}$

Parameters: A functionality $\mathcal{F}_{\text{AQB}}$ with security parameter m, basis set $\mathcal{B}$.
Parties: The sender Alice and the receiver Bob.
Setup phase

1) Alice chooses a time index τ and sends it to Bob. Then, Alice and Bob record the outputs $(\mathcal{C}(s) \oplus \ell, j, \tau)$ and $(\ell, |e_j^s\rangle, \tau)$, respectively, sent by $\mathcal{F}_{\text{AQB}}$.

Commit phase
Inputs: Alice gets a bit b.

2) Alice commits the bit b to Bob by sending him $c = b \oplus j$.
3) After receiving c from Alice, Bob outputs the message *receipt*.

Open phase
Inputs: Alice gets the message *open*.

4) Alice sends $q = \mathcal{C}(s) \oplus \ell$ to Bob.
5) If $q \oplus \ell \notin \text{Im}(\mathcal{C})$, Bob outputs *error*. Otherwise, he computes $s = \mathcal{C}^{-1}(q \oplus \ell)$, measures his qubit $|e_j^s\rangle$ in the basis $(|e_0^s\rangle, |e_1^s\rangle)$ to obtain j, and outputs $b = c \oplus j$.

Figure 3: Bit commitment protocol $\pi_{\text{BC}}^{\text{AQB}}$ using the functionality $\mathcal{F}_{\text{AQB}}$.

4.1 Security analysis

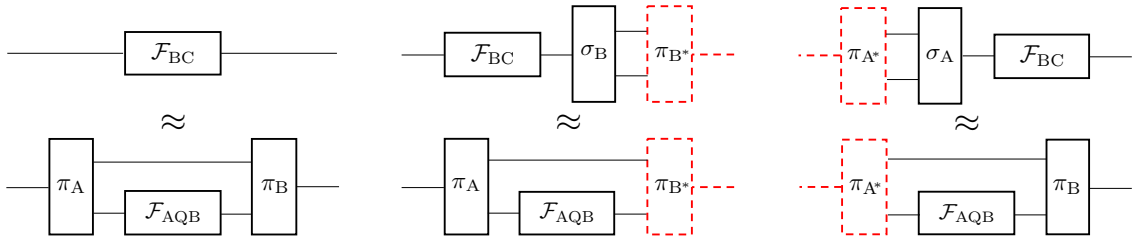
We proceed now to prove the security of $\pi_{\text{BC}}^{\text{AQB}}$ in the UC framework. Figure 4 shows diagrams depicting the equivalences that need to be satisfied in order for the protocol to be UC secure. In Fig. 4(b) and 4(c), the boxes $\sigma_{\text{B/A}}$ represent the programs that convert protocol adversaries $\pi_{\text{B}^*/\text{A}^*}$ into simulated adversaries $\sigma_{\text{B/A}}\pi_{\text{B}^*/A^*}$ interacting with the ideal functionality $\mathcal{F}_{\text{BC}}$.

1) *Soundness:* Because the value of j is fixed by $\mathcal{F}_{\text{AQB}}$ during the setup phase, when Alice sends $c = b \oplus j$, she fixes her commitment to b. Measuring $|e_j^i\rangle$ in the $\mathcal{B}_i$ basis is guaranteed to output j. Therefore, Bob always obtains b by taking $c \oplus j$.
2) *Concealing:* Consider the following program for the converter σ_B. During the setup phase, σ_B simulates $\mathcal{F}_{\text{AQB}}$ to generate the values $s, \mathcal{C}(s), \ell, j$ and the qubit $|e_j^i\rangle$, then sends $(\ell, |e_j^i\rangle)$ to the adversary. In the commit phase, upon receiving (*commit*) from $\mathcal{F}_{\text{BC}}$, it sends $c' = j$ to the adversary. During the open phase, after receiving (*open, b*) from $\mathcal{F}_{\text{BC}}$, if $b = 0$ it sends $q = \mathcal{C}(s) \oplus \ell$ to the adversary, otherwise $q' = \mathcal{C}(\tilde{s}) \oplus \ell$. The concealing property comes by noting that the behaviour of σ_B is the same regardless of the bit that was sent to $\mathcal{F}_{\text{AQB}}$, hence there is no algorithm π_{B^*} that can guess the committed bit with probability greater than $1/2$.
3) *Binding:* We show that, whenever the protocol does not output error, the equivalence in Fig. 4(c) holds, except with negligible probability. Consider the following program for the converter σ_A. During the setup phase, σ_A simulates $\mathcal{F}_{\text{AQB}}$ to generate the values $s, \mathcal{C}(s), \ell, j$ and the qubit $|e_j^i\rangle$. It then sends $(q = \mathcal{C}(s) \oplus \ell, j)$ to the adversary. During the commit phase, upon receiving a bit c' from the adversary, it computes $b = c' \oplus j$ and outputs (*commit, b*) to $\mathcal{F}_{\text{BC}}$. During the opening phase, upon receiving q' from the adversary, if $q' = q$ it outputs (*open*) to $\mathcal{F}_{\text{BC}}$.
   Bob outputs error whenever the string q' sent by Alice is such that $q' \oplus \ell \notin \text{Im}(\mathcal{C})$. From the soundness property we know that when $q = q'$ Bob opens $b = c' \oplus j$. We are interested in the case when $q' \neq q$ (in which case the upper part of Fig. 4(c) does not open anything) and $q' \oplus \ell \in \text{Im}(\mathcal{C})$ (in which case the bottom part opens to a bit). Because $f(s)$ is injective and perfectly non-linear, finding such q' given q is equivalent to finding s given $\mathcal{C}(s) \oplus \ell$, which the adversary can do only with negligible probability in m.

5. DISCUSION AND CONCLUSIONS

In this paper, we presented the AQB model, in which is possible to construct a perfectly secure Bit-commitment protocol. In this model, a trusted party T that does not get any input and only provides split information to the other parties is assumed. We say that the AQB model represents a minimal assumption for composable BC security because T need not be online during the commit and opening phases and does not get information about the committed bit. Moreover, all of the algorithms run by the parties have complexity linear in the security parameter m. Similarly, the total amount of information exchange between the parties is exactly $m + 1$ bits,

Authorized licensed use limited to: University of Exeter. Downloaded on June 14,2020 at 15:21:39 UTC from IEEE Xplore. Restrictions apply.



(a) Soundness (b) Concealing (c) Binding

Figure 4: Diagram (a) corresponds to the soundness (correctness) property by showing the equivalence between the ideal commitment functionality $\mathcal{F}_{\mathrm{COM}}$ and the protocol for honest parties. Diagrams (b) and (c) correspond to security against dishonest Bob and Alice, respectively. The dashed boxes represent the dishonest party running a different algorithm $\pi_{\mathrm{A}^*/\mathrm{B}^*}$.

except between T and Bob, where one bit is replaced by a qubit sent additionally over a quantum channel. It is worth noting though, that the communication between Alice and Bob is classical and authenticated.

In a practical implementation scenario, finding an appropriate perfectly (or almost perfectly) non-linear bijection is a non-trivial task. Nevertheless, there are well known functions (such as the AES block cipher), which are known to be highly non-linear and very efficient to compute, that can replace them for practical security requirements.

ACKNOWLEDGEMENTS

The authors acknowledge the support of SQIG – Security and Quantum Information Group, under the project UID/EEA/50008/2019, the project QuantumMining POCI-01-0145-FEDER-031826 funded by FCT through national funds, by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme (COMPETE 2020), and by Regional Operational Program of Lisbon, and the project Predict PTDC/CCI-CIF/29877/2017, funded by FCT. ML thanks the support of DP-PMI and FCT grant PD/BD/114334/2016. PY thanks the support of DP-PMI and FCT grant PD/BD/113648/2015. NP thanks FCT grant CEECIND/04594/2017. AS also acknowledges fundings of LASIGE under the project ref. UID/CEC/00408/2013 of FCT. The authors also acknowledge the projects Confident PTDC/EEI-CTP/4503/2014.

REFERENCES

[1] M. Blum, "Coin flipping by telephone a protocol for solving impossible problems," *SIGACT News*, vol. 15, no. 1, pp. 23–27, 1983. [Online]. Available: http://doi.acm.org/10.1145/1008908.1008911

[2] A. Almeida, R. Loura, N. Paunković, N. Silva, N. Muga, P. Mateus, P. André, and A. Pinto, "A brief review on quantum bit commitment," vol. 9286, 08 2014, p. 92861C.

[3] A. C.-C. Yao, "Security of quantum protocols against coherent measurements," in *Proceedings of the Twenty-seventh Annual ACM Symposium on Theory of Computing*, ser. STOC '95. New York, NY, USA: ACM, 1995, pp. 67–75.

[4] O. Goldreich, S. Micali, and A. Wigderson, "How to play any mental game," in *Proceedings of the nineteenth annual ACM symposium on Theory of computing*. ACM, 1987, pp. 218–229.

[5] R. Canetti and M. Fischlin, "Universally composable commitments," in *Annual International Cryptology Conference*. Springer, 2001, pp. 19–40.

[6] D. Mayers, "Unconditional security in quantum cryptography," *J. ACM*, vol. 48, no. 3, pp. 351–406, 2001. [Online]. Available: http://doi.acm.org/10.1145/382780.382781

[7] H. K. Lo and H. F. Chau, "Is quantum bit commitment really possible?" *Phys. Rev. Lett.*, vol. 78, pp. 3410–3413, 1997. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.78.3410

[8] R. Canetti, "Universally composable security: A new paradigm for cryptographic protocols," in *Proceedings of the 42Nd IEEE Symposium on Foundations of Computer Science*, ser. FOCS '01. Washington, DC, USA: IEEE Computer Society, 2001, pp. 136–. [Online]. Available: http://dl.acm.org/citation.cfm?id=874063.875553

[9] R. Dowsley, J. v. Graaf, J. Müller-Quade, and A. C. Nascimento, "On the composability of statistically secure bit commitments," *Journal of Internet Technology*, vol. 14, no. 3, p. 509, 2013.

[10] R. Dowsley, J. Müller-Quade, and A. Nascimento, "On possibility of universally composable commitments based on noisy channels," *Anais do VIII Simpósio Brasileiro em Segurança da Informaçao e de Sistemas Computacionais, SBSEG*, pp. 103–114, 2008.

Authorized licensed use limited to: University of Exeter. Downloaded on June 14,2020 at 15:21:39 UTC from IEEE Xplore. Restrictions apply.

Journal of Physics:
Conference Series

PURPOSE-LED PUBLISHING™

PAPER • **OPEN ACCESS**

Gravity-matter entanglement in Regge quantum gravity

To cite this article: Nikola Paunković and Marko Vojinović 2016 *J. Phys.: Conf. Ser.* **701** 012035

View the article online for updates and enhancements.

You may also like

- A generalized Hartle–Hawking wave function
Stephon Alexander, Gabriel Herczeg and João Magueijo
- Gauge protected entanglement between gravity and matter
N Paunkovi and M Vojinovi
- A little excitement across the horizon
Keith K Ng, Chen Zhang, Jorma Louko et al.



This content was downloaded from IP address 147.91.81.84 on 19/09/2025 at 14:45

Journal of Physics: Conference Series **701** (2016) 012035 doi:10.1088/1742-6596/701/1/012035

Gravity-matter entanglement in Regge quantum gravity

Nikola Paunković[1] and Marko Vojinović[2]

[1]SQIG — Security and Quantum Information Group, Instituto de Telecomunicaõoes, and Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1049-001, Lisboa, Portugal

[2]Grupo de Física Matemática, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Edifício C6, 1749-016 Lisboa, Portugal

E-mail: npaunkov@math.ist.utl.pt, vmarko@ipb.ac.rs

Abstract.

We argue that Hartle–Hawking states in the Regge quantum gravity model generically contain non-trivial entanglement between gravity and matter fields. Generic impossibility to talk about "matter in a point of space" is in line with the idea of an emergent spacetime, and as such could be taken as a possible candidate for a criterion for a plausible theory of quantum gravity. Finally, this new entanglement could be seen as an additional "effective interaction", which could possibly bring corrections to the weak equivalence principle.

Introduction. The unsolved problems of interpreting quantum mechanics (QM) and formulating quantum theory of gravity (QG) are arguably the two most prominent ones of the twentieth century theoretical physics. Up to date, most of the efforts to solve the two were taken independently. Indeed, the majority of the interpretations of QM do not involve explicit dynamical effects (with notable exceptions of the spontaneous collapse and the de Broglie–Bohm theories), while the researchers from the QG community often adopt the many-world interpretation of QM. Nevertheless, the two problems share a number of similar unsolved questions and counter-intuitive features, such as nonlocality: entanglement-based quantum nonlocality, as well as the anticipated explicit dynamical nonlocality in QG (a consequence of quantum superpositions of different gravitational fields, i.e., different spacetimes and their respective causal orders). We analyse the generic entanglement between gravitational and matter fields in the Regge model of quantum gravity, and its possible impact to the fundamental questions regarding QM and QG.

Regge quantum gravity model. A simple toy model of quantum gravity with matter fields is the Regge quantum gravity with one real scalar field, whose construction can be motivated by the Loop Quantum Gravity research program [1, 2]. The path integral of the model is

$$Z_T = \int \mathcal{D}L \int \mathcal{D}\Phi \;\; \exp\left[iS_{\text{Regge}}(L) + iS_{\text{matter}}(L,\Phi)\right], \tag{1}$$

where L are the lengths of the edges of the triangulation T of a 4-manifold $\mathcal{M}_4$, and Φ are the values of the scalar field in 4-simplices of T. The measure terms $\mathcal{D}L$ and $\mathcal{D}\Phi$ are defined via discretization induced by T. The actions S_{Regge} and S_{matter} represent lattice discretizations


Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

of the Einstein–Hilbert action for gravity and an action of the scalar field coupled to gravity, respectively. See [3, 4] for details.

A generic state of the system, belonging to the kinematical Hilbert space $\mathcal{H}_G \otimes \mathcal{H}_M$, is

$$|\Psi\rangle = \int \mathcal{D}l \int \mathcal{D}\phi \;\; \Psi(l,\phi) \;\; |l\rangle|\phi\rangle \,. \tag{2}$$

However, since gravity is a theory with constraints, not every kinematical state is allowed, so we must choose the coefficients $\Psi(l,\phi)$ such that $|\Psi\rangle$ is an element of the physical Hilbert space $\mathcal{H}_{\text{phys}} \subset \mathcal{H}_G \otimes \mathcal{H}_M$. One such class of states are the Hartle–Hawking (HH) states [5], defined by the following choice of the coefficients, for a given triangulation T:

$$\Psi(l,\phi) = \Psi_{\text{HH}}(l,\phi) \equiv \int \mathcal{D}L \int \mathcal{D}\Phi \;\; \exp\Big[iS_{\text{Regge}}(L,l) + iS_{\text{matter}}(L,\Phi,l,\phi)\Big] \,. \tag{3}$$

This expression differs from (1) in that the triangulation T is now assumed to have a nontrivial 3-dimensional boundary ∂T, and that the variables l, ϕ living on the boundary are not integrated over, in contrast to the bulk variables L and Φ.

Using (2) and (3), one can calculate the reduced density matrix of the Hartle–Hawking state,

$$\hat{\rho}_M = \text{Tr}_G \, |\Psi\rangle\langle\Psi| = \int \mathcal{D}\phi \int \mathcal{D}\phi' \;\; \left[\int \mathcal{D}l \;\; \Psi_{\text{HH}}(l,\phi)\Psi^*_{\text{HH}}(l,\phi')\right] \;\; |\phi\rangle\langle\phi'| \,,$$

where the integral in the brackets can be denoted as $Z_{T\cup\bar{T}}(\phi,\phi')$. The resulting density matrix can then be tested for entanglement by checking if the trace of its square equals one [6]. In the Regge quantum gravity model the path integrals reduce to a finite number of ordinary integrals, which can then in principle be evaluated. For a generic triangulation, we obtain

$$\text{Tr}_M \, \hat{\rho}_M^2 = \int \mathcal{D}\phi \int \mathcal{D}\phi' \;\; \left|Z_{T\cup\bar{T}}(\phi,\phi')\right|^2 \neq 1 \,,$$

i.e., the gravitational and scalar degrees of freedom in the generic HH state are entangled.

Discussion. In [7] Penrose argues that gravity-matter entanglement is at odds with (classical) spacetime, seen as a (four-dimensional) differentiable manifold. In light of this, our result could be seen as a quantitative indicator that in quantum gravity one cannot talk of "matter in a point of space", i.e., this result could be seen as a confirmation of a "spacetime as an emergent phenomenon".

Thus, generic gravity-matter entanglement could be seen as a possible candidate for a criterion for a plausible theory of quantum gravity.

Entanglement is in standard quantum mechanics a generic consequence of the interaction. This new entanglement can be regarded as a consequence of an effective interaction (such as the "exchange interactions", which are a consequence of quantum statistics). This additional "effective interaction" can potentially lead to corrections to the weak equivalence principle.

Acknowledgments
NP was partially supported, under the CV-Quantum internal project at IT, by FCT PEst-OE/EEI/LA0008/2013 and UID/EEA/50008/2013. MV was supported by the FCT grant SFRH/BPD/46376/2008, the FCT project PEst-OE/MAT/UI0208/2013, and partially by the project ON171031 of the Ministry of Education, Science and Technological Development, Serbia.

References

[1] Rovelli C 2004 *Quantum Gravity* (Cambridge: Cambridge University Press)
[2] Rovelli C and Vidotto F 2014 *Covariant Loop Quantum Gravity* (Cambridge: Cambridge University Press)
[3] Miković A and Vojinović M 2012 *Class. Quant. Grav.* **29** 165003
[4] Miković A 2013 *Rev. Math. Phys.* **25** 1343008
[5] Hartle J B and Hawking S W 1983 *Phys. Rev. D* **28** 2960
[6] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[7] Penrose R 1996 *Gen. Relativ. Gravit.* **28** 581

Enhancing Privacy with Quantum Networks

Paulo Mateus, Nikola Paunković, João Rodrigues, and André Souto

SQIG- Instituto de Telecomunicações and
DM - Instituto Superior Técnico - Universidade de Lisboa, Portugal

Abstract. Using quantum networks to distribute symmetric keys has become a usable and commercial technology available under limitations that are acceptable in many application scenarios. The fact that the security is implemented directly at the hardware level, and moreover, relies on the laws of physics instead of conjectured hardness assumptions, justifies the use of quantum security in many cases. Limitations include 100 km communication range and installation of quantum channels between each pair of users of the network. Presently, with the current lack of trust in commercial security solutions, mostly due to the Snowden crisis, there is the need to improve such solutions. In this paper we discuss how quantum networks can be used to setup secure multiparty computation (SMC), allowing for instance for private data mining, electronic elections among other security functionalities. SMC relies mostly on establishing an efficient oblivious transfer protocol. We present a bit-string quantum oblivious transfer protocol based on single-qubit rotations that can be implemented with current technology based on optics and whose security relies only on the laws of physics.

1 Introduction

Security is the most important factor for building trust and confidence between consumers/population and companies/State; this trust has been severely damaged with many recent events such as the "Snowden crisis" and the Open SSL critical bug, and as such, private companies and state providers are pressured to improve the security of their products. In this paper we discuss how quantum security protocols can be integrated in a classical setting to provide multiparty-secure computation.

Two seminal works have driven most of the research in the area quantum security: the quantum polynomial time factorization algorithm proposed by Shor [7]; and the quantum public key agreement protocol BB84, proposed by Bennett and Brassard [1]. While Shor's algorithm raises the threat of making widely used cryptographic systems (via classic communication channels) completely obsolete by a breakthrough in quantum hardware, the BB84 protocol shows that quantum communication channels allow public perfect security in the context of an authenticated channel.

Due to Shor's factoring algorithm, research on (asymmetric) cryptography shifted significantly. Presently, one of the most important problems in the area is to find one-way functions robust to quantum attacks. Indeed, Shor's algorithm

B. De Decker and A. Zúquete (Eds.): CMS 2014, LNCS 8735, pp. 147–153, 2014.
© IFIP International Federation for Information Processing 2014

is able to attack all cryptosystems based on factorization and discrete logarithm, even in the elliptic curve setting, which accounts for essentially everything that is used in practice and is based on asymmetric keys.

On the other hand, BB84 is already commercially available through peer-to-peer optical networks. It is worth pointing out that quantum channels sending an arbitrarily amount of quantum information can already be produced using cheap technology. Moreover, much research is being done to develop quantum networks and routers using traditional optical fibers and laser satellite communications. It is expected that quantum networks will be available much sooner than quantum computers and thus, it is fundamental to understand which security and distributed protocols can benefit from quantum technology.

Secure multiparty computation is an abstraction of many security functionalities, including private data mining, e-voting, verifiable secret sharing, verifiable computing, among others. In general terms, the goal of secure multiparty computation among n parties is to compute a function of n secret inputs, one for each party, such that at the end of the computation the output of the function is known to all parties, while keeping the inputs secret.

It is well known that to setup secure multiparty computation it is enough to establish oblivious transfer (OT) protocol between two-parties using Yao's garbled circuits [8] (see a more modern discussion in [3]). The first OT protocol was presented by Rabin [6] and its security relies on the hardness assumption that factoring large integers is difficult in polynomial time. OT can be seen as a game played by two parties, Alice and Bob. Alice wants to share a number of secret messages with Bob such that, on average, Bob receives half of those messages (the protocol is *concealing*), while keeping Alice unaware to which messages Bob got (the protocol is *oblivious*). A protocol achieving these properties is called an OT Protocol. An OT protocol is made out of two parts: the transferring phase and the opening phase. In the former Alice sends the secret message to Bob; in the latter Alice unveil enough information that allows Bob to recover the secret with probability 1/2.

The main contribution of this paper is to propose an OT protocol that can be implemented over quantum optical networks using currently available technology. Such OT protocol can then be used to establish a secure multiparty computation using classical infrastructure. We introduce a quantum oblivious transfer protocol for bit-strings, based on the recently proposed public key crypto-system in [5]. Each bit of the string to be transferred is encoded in a qubit (quantum bit), a particular quantum state, in such a way that states corresponding to bit-values 0 and 1, respectively, form an orthonormal basis. The key point of the protocol is that for each qubit, the encoding basis is chosen at random, from some discrete set of bases.

Next section provides a brief survey of quantum information, including basic definitions and important results necessary for understanding our proposal. Section 3 describes our proposal for a bit-string oblivious transfer protocol and discusses its correctness and security. Finally, we summarize the results and discuss future directions of research.

2 Preliminaries

In this section, we provide notation, necessary definitions and results for defining and reasoning about the security of our proposal.

For a complete study of quantum information we suggest the reading of [4]. Here we present some relevant notions. According to the postulates of quantum mechanics, the state of a closed quantum system is represented by a unit vector from a complex Hilbert space $\mathcal{H}$, and its evolution is described by a unitary transformation on $\mathcal{H}$. In this paper we work only with finite-dimensional Hilbert spaces reflecting the realistic examples of systems with finite number degrees of freedom (strings of quantum bits, i.e. qubits).

Contrarily to the classical case where a bit can only have values 0 or 1, in the quantum case, a *qubit* can be in a unit superposition of 0 or 1 denoted by $\alpha|0\rangle + \beta|1\rangle$ with complex coefficients α and β such that $|\alpha|^2 + |\beta|^2 = 1$. The Dirac notation $|0\rangle$ and $|1\rangle$ denotes vectors forming an orthonormal basis of a 2-dimensional complex vector space. Note that we can define many orthonormal bases for that space, such as $\left\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right\}$, but it is common to distinguish the basis $\{|0\rangle, |1\rangle\}$ from all the others, and call it the *computational basis*.

The state of two qubits is from the tensor product of single-qubit spaces, that is,

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

with $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. The state $|\psi\rangle$ is said to be *separable* if

$$|\psi\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha'|0\rangle + \beta'|1\rangle) = \alpha\alpha'|00\rangle + \alpha\beta'|01\rangle + \alpha'\beta|10\rangle + \beta\beta'|11\rangle .$$

Otherwise, it is called *entangled*. Although entangled states are particularly important in quantum information, in this paper we only work with separable states. Note that a system with k qubits can be described by a unit vector over a space with dimension 2^k.

One of the most important results of quantum information states that the maximal information that can be stored in a qubit is the same as that contained in a bit. This means that we cannot extract more than a bit of information from a qubit, although there is potentially an infinite number of states available to encode in a qubit. The reason for this is that it is impossible to obtain coefficients α and β from a single qubit in a state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Indeed, what is possible is to perform a measurement given by an orthogonal decomposition of the Hilbert space $\mathcal{H} = \bigoplus_{i=1}^{d} \mathcal{H}_i$, with P_i being the projectors onto $\mathcal{H}_i$. Then, upon performing such a measurement on a qubit in state $|\psi\rangle \in \mathcal{H}$, there are d possible outcomes $\{1, \ldots, d\}$, where the probability of observing $i \in \{1, \ldots, d\}$ is given by $\|P_i|\psi\rangle\|$, and then the state evolves to $P_i|\psi\rangle / \|P_i|\psi\rangle\|$. For instance, the outcome of a measurement of a qubit can only take two possible values.

To understand the protocol we need to consider a function that is easy to compute, but, without the help of a secret trapdoor, it is impossible to invert with

non-negligible probability according to the laws of quantum physics. One candidate for such a function was proposed in [5] which uses sinlge-qubit rotations and is given by

$$f(s) = R(s\theta_n)\,|0\rangle = \cos(s\theta_n/2)\,|0\rangle + \sin(s\theta_n/2)\,|1\rangle$$

where, for some fixed n, $s \in \{0, \ldots, 2^n - 1\}$, $\theta_n = \pi/2^{n-1}$ and $\{|0\rangle, |1\rangle\}$ is a fixed computational basis (i.e., f is not a function of a quantum state). Moreover, f can be used to construct a quantum trapdoor function $F(s, b)$, where s is the trapdoor information for learning an unknown bit b [5]:

$$F(s, b) = R(b\pi)f(s) = R(b\pi)R(s\theta_n)\,|0\rangle = R(s\theta_n + b\pi)\,|0\rangle\,.$$

Note that inverting F (learning both s and b) is at least as hard as inverting f. In [5] it was shown that every binary measurement that could be used to infer unknown bit b would outcome a completely random value. Nevertheless, if s is known, by applying the rotation $R(-s\theta_n)$ to $F(s, b)$ and measuring the result in the computational basis, one obtains b with certainty.

Using the properties of f and F a secure public key cryptographic protocol was proposed in [5]: using the private key s, the public key is generated by computing $f(s)$; the encryption of a secret message corresponds to computing $F(s, b)$; and the decryption of the message corresponds to inversion of $F(s, b)$, using the trapdoor information s.

Finally, in order to guarantee that at the end of the OT protocol Bob knows if he got the message m or not, Alice is required to send both m and $h(m)$, where h is a *universal hash function*. A hash function maps strings to other strings of smaller size . Bellow, we present a definition of universal hash function and a basic result.

Definition 1. *Consider two sets A and B of size a and b, respectively, such that $a > b$, and consider a collection $\mathbb{H}$ of hash functions $h : A \to B$. If*

$$\Pr_{h \in \mathbb{H}}[h(x) = h(y)] \leq \frac{1}{b}$$

then $\mathcal{H}$ is called a universal family of hash functions.

Theorem 1. *Let $\mathbb{H}$ be a collection of hash functions $h : A \to B$, where A and B are sets of size a and b, respectively, such that $a > b$. The size of a set A_x of strings $x \in A$ mapped to the same hash value $h(x)$ is at most N/b.*

In our particular case we consider A and B as the sets of strings of length ℓ and $\ell/2$ respectively. Hence, there are $2^{\ell/2}$ different strings for each hash value (for an overview see [2]).

3 Oblivious Transfer

Having set the required definitions and results, our protocol works as follows:

Protocol 1 (Oblivious transfer)

Message to transfer $m = m_1 \dots m_k$;
Security parameter $n, \theta_n = \pi/2^{n-1}$ *and a hash function* $h : \{0,1\}^k \to \{0,1\}^{k/2}$;
Secret key $s = (s_1, \dots, s_{3k/2})$, *where each* $s_i \in \{0, \dots, 2^n - 1\}$.

Transfer Phase

1. *Alice chooses uniformly at random a bit* $a \in \{0,1\}$ *and prepares the following state:*

$$|\psi\rangle = \bigotimes_{i=1}^{k} R(m_i\pi + (-1)^a \times s_i\theta_n)\,|0\rangle \bigotimes_{i=1}^{k/2} R(h_i(m)\pi + (-1)^a \times s_{i+k}\theta_n)\,|0\rangle$$

   (Note that $h_i(m)$ *represents the* i^{th} *bit of the binary string* $h(m)$*).*
2. *Alice sends the state* $|\psi\rangle$ *to Bob.*

Opening Phase

3. *Alice sends* $s = (s_1, \dots, s_{3k/2})$ *and* n *to Bob.*
4. *Bob checks if* s *is likely to be a possible output of a random process by performing a statistical test.*
5. *Bob chooses uniformly at random* $a' \in \{0,1\}$ *and applies* $R((-1)^{a'} s_i\theta_n)$ *to each qubit of* $|\psi\rangle$.
6. *Bob applies the measurement operator* $M = (0 \times |0\rangle\langle 0| + 1 \times |1\rangle\langle 1|)^{\otimes 3k/2}$.
7. *Let* $m' \cdot h'$ *be the message that Bob recovers. He checks if* $h' = h(m')$. *If that is the case then Bob is almost sure that* $m' = m$, *otherwise he knows that* m' *is not the correct message.*

In the following, we discuss the security of our oblivious transfer protocol, showing that: if both agents are honest, Bob will obtain the message m with probability $1/2$ (the protocol is sound); if Alice plays fair, Bob is not able to recover m before the opening phase (the protocol is concealing); if Bob is honest, then Alice is unaware if Bob got m or not (the protocol is oblivious).

To state the results we need the notion of negligible function. $\varepsilon : \mathbb{N} \to \mathbb{R}$, a nonnegative function is called *negligible* if for every polynomial p and sufficiently large k we have $\varepsilon(k) \leq 1/p(k)$.

First, we provide the reasoning for the soundness of our protocol.

Theorem 2. *If both parties are honest, then with probability* $1/2 + \varepsilon(k)$ *Bob will get the right message, where* $\varepsilon(k)$ *is negligible function on the size of the message* $m = m_1 \dots m_k$.

Notice that if Alice and Bob are honest then the choice of rotation direction of both will differ with probability 1/2. Only when they are different, i.e., Bob undo Alice's rotation and obtains the states in computational bases, Bob is ensured to recover the message. When Bob rotates in the same direction of Alice, the results of Bob's measurement are random and hence the probability of recovering m in this case is a negligible function on the size of the message m.

We proceed by discussing the concealing property of the protocol.

Theorem 3. *If Alice is honest, the probability of Bob recovering Alice's message before the opening phase is negligible. Furthermore, after the opening phase Bob recovers the message, up to a negligible value, with probability* 1/2.

The first part of the theorem follows directly from the security of the public key encryption schemes presented in in [5]: without knowing the secret key s and the rotation direction a, Bob's description of a message m is given by a completely mixed state. The second part follows from a similar argument to the previous theorem.

To finish the security discussion we argue that the protocol is unconditionally oblivious.

Theorem 4. *The Protocol 1 is oblivious, i.e., at the end of the protocol Alice does not know whether Bob received the right message of not.*

During the execution of the protocol, there is no information traveling from Bob to Alice. Therefore, in order to increase the probability of learning if Bob received the message m or not, Alice has to perform the following cheating strategy: instead of sending $|\psi\rangle$, Alice sends a cheating state $|\psi_{ch}\rangle$ for which Bob will open the desired message with a probability greater than 1/2 (possibly with certainty). This is impossible, unless with negligible increase $\varepsilon(l)$, bounded above by $\frac{1}{2}\left(1+\cos^{2l}(\pi/8)\right)$, where l is the number of s_i's for which $s_i\theta_n \in [\pi/8; 3\pi/8]$.[1]

4 Conclusions

In this paper we proposed a scheme for oblivious transfer of a bit-string message. Its security is based on laws of quantum physics. We reasoned about its security and showed that the protocol is unconditionally oblivious and concealing. Our protocol can be implemented with today's technology using optical equipment. Moreover, the protocol can be integrated with existing classical networks to achieve secure multiparty computation, and promote an extra level of security on such functionality.

Using single-qubit rotations have been proved useful in designing quantum security protocols, such as the presented oblivious transfer and the previously proposed public key cryptographic scheme [5]. This opens a number of possible future applications of single-qubit rotations in designing several secure protocols such as quantum bit-string commitment protocol and undeniable signatures.

[1] Since the values of s_i's are required to be random, the expected value of l is $k/4$, with the standard deviation $\sigma = \sqrt{k}/4$.

Acknowledgments. This work was partially supported by FCT projects ComFormCrypt PTDC/EIA-CCO/113033/2009 and PEst-OE/EEI/LA0008/2013, and SQIG's initiative PQDR (Probabilistic, Quantum and Differential Reasoning). João Rodrigues and André Souto also acknowledges respectively the FCT grants SFRH / BD / 75085 / 2010 and SFRH / BPD / 76231 / 2011.

References

1. Bennett, C., Brassard, G.: Quantum Cryptography: Public Key Distribution and Coin Tossing. In: Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, pp. 175–179. IEEE Press, New York (1984)
2. Carter, J., Wegman, M.: Universal classes of hash functions. J. Comput. Syst. Sci. 18(2), 143–154 (1979)
3. Lindell, Y., Zarosim, H.: On the feasibility of extending oblivious transfer. In: Sahai, A. (ed.) TCC 2013. LNCS, vol. 7785, pp. 519–538. Springer, Heidelberg (2013)
4. Nielsen, M., Chuang, I.: Quantum Computation and Quantum Information, 1st edn. Cambridge University Press (January 2004)
5. Nikolopoulos, G.: Applications of single-qubit rotations in quantum public-key cryptography. Phys. Rev. A 77, 032348(2008)
6. Rabin, M.: How to exchange secrets by oblivious transfer (1981)
7. Shor, P.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM J. Comput. 26(5), 1484–1509 (1997)
8. Yao, A.C.-C.: How to generate and exchange secrets. In: Proceedings of the 27th Annual Symposium on Foundations of Computer Science, SFCS 1986, pp. 162–167. IEEE Computer Society, Washington, DC (1986)

A Brief Review on Quantum Bit Commitment

Álvaro J. Almeida[a,b], Ricardo Loura[c,d], Nikola Paunković[c,d], Nuno A. Silva[b,e], Nelson J. Muga[b,e], Paulo Mateus[c,d], Paulo S. André[b,f], Armando N. Pinto[b,e]

[a]Department of Physics, University of Aveiro, Aveiro, Portugal
[b]Instituto de Telecomunicações, Aveiro, Portugal
[c]SQIG-Instituto de Telecomunicações, IST-TU Lisbon, Lisboa, Portugal
[d]Department of Mathematics, IST, Technical University of Lisbon, Lisboa, Portugal
[e]Department of Electronics, Telecommunications and Informatics, University of Aveiro, Aveiro, Portugal
[f]Department of Electrical and Computer Engineering, Instituto Superior Técnico, Lisboa, Portugal

ABSTRACT

In classical cryptography, the bit commitment scheme is one of the most important primitives. We review the state of the art of bit commitment protocols, emphasizing its main achievements and applications. Next, we present a practical quantum bit commitment scheme, whose security relies on current technological limitations, such as the lack of long-term stable quantum memories. We demonstrate the feasibility of our practical quantum bit commitment protocol and that it can be securely implemented with nowadays technology.

Keywords: Quantum Bit Commitment, Quantum Cryptography.

1. INTRODUCTION

Bit commitment is a fundamental primitive in cryptography which allows two untrustful parties to do a secure commitment. It consists in two phases: *commitment* and *revealing*. In the commitment phase, the one that commits, Alice, has to choose between a bit value, 0 or 1. After commitment, Alice gives her choice to Bob during the revealing phase. In order to be secure, the protocol must be both binding and concealing. To be binding means that Alice cannot change her choice later in time, in particular during the revealing phase. To be concealing means that Bob cannot learn Alice's commitment before she reveals it to him.

In classical cryptography, the security of bit commitment lies in assumptions of computational complexity such as the difficulty of factoring large numbers.[1] However, with the increase of the computational power and the rise of quantum computers, its security can be broken.[2] In fact, it can be proven that unconditionally secure bit commitment based only on classical resources is impossible.[3] It was also demonstrated that even if Alice and Bob have access to quantum resources, unconditionality secure quantum bit commitment (QBC) is impossible within nonrelativistic physics.[4,5] As in classical cryptography, QBC schemes must rely on physical assumptions, e.g. the quantum memory that an attacker can use is noisy.[6] The situation changed when it was found that using Minkowsky causality in a relativistic scenario it is possible to build unconditionally secure QBC protocols.[7] This sort of protocols are starting to be implemented nowadays.[3,8]

Further author information: (Send correspondence to Álvaro J. Almeida)
Álvaro J. Almeida: E-mail: aalmeida@av.it.pt, Telephone: +351 234 377 900
Ricardo Loura: E-mail: ricardoloura@gmail.com
Nikola Paunkovic: E-mail: npaunkov@math.ist.utl.pt
Nuno A. Silva: E-mail: nasilva@av.it.pt
Nelson J. Muga: E-mail: muga@av.it.pt
Paulo Mateus: E-mail: pmath@math.ist.utl.pt
Paulo S. André: E-mail: paulo.andre@lx.it.pt
Armando N. Pinto: E-mail: anp@ua.pt

Second International Conference on Applications of Optics and Photonics, edited by Manuel Filipe P. C. Martins Costa, Rogério Nunes Nogueira, Proc. of SPIE Vol. 9286, 92861C · © 2014 SPIE · CCC code: 0277-786X/14/$18 · doi: 10.1117/12.2063733

We review the state of the art of bit commitment schemes, from classical to quantum, and present the main achievements and drawbacks. A practical QBC scheme is also presented and experimentally verified in a back to back case scenario. Finally we present the main conclusions of this work.

2. FROM CLASSICAL BIT COMMITMENT TO UNCONDITIONALLY SECURE RELATIVISTIC QUANTUM BIT COMMITMENT

2.1 Review of the State of the Art

The importance of encryption and secrecy in communication is well studied in the mathematical theory of communication from Shannon.[9] In this influential article, Shannon presented what he considered the essential five parts of a communication system: the information source, that produces a message; the transmitter which operates on the message to create a signal that can be transmitted through a channel; the channel which is the medium of transmission of the signal; the receiver which performs the inverse operation done by the transmitter in order to reconstruct the message and the destination, person or machine, to whom or which the message is intended. Due to the need for security in new applications such as electronic bank transactions or the improvement of computer technology, new encryption schemes were created, like the *data encryption standard* (DES)[10] or the *advanced encryption standard* (AES).[11] Several years before the proposed schemes, Stephen Wiesner had the idea to make quantum banknotes, that was not accepted immediately, giving rise to a complete new level of security in what would be called quantum cryptography.[12]

In 1981, Blum presented the notion of bit commitment in his work about coin flipping over telephone.[13] The concept of commitment would be formalized by Brassard et al. only in 1988.[14] In fact, the first quantum cryptographic protocol, which was proposed in 1984 by Bennett and Brassard, widely known as BB84,[15] was already a preliminary version of a bit commitment protocol, since the idea was to associate the commitments of 0 and 1 with two complementary observables. In a later paper, a QBC protocol, known as BCJL, was claimed by the authors to be unconditionally secure.[16] In that work, the authors presented an argument in which, according to the laws of quantum physics, the participants in the protocol are only allowed to cheat with an arbitrarily small probability. Despite the initial optimism to achieve an unconditionally secure QBC protocol, Mayers[4] and independently Lo and Chau,[5] proved a no-go theorem showing that unconditionally secure QBC is impossible unless relativistic effects are used. This impossibility comes from the fact that a cheating strategy using EPR pairs can always be implemented. Thus, different approaches have been presented in order to avoid the no-go theorem.[17] These proposals are based on relaxing some of the assumptions, such as using noisy/bounded quantum memories.[6] Unfortunately none of these attempts achieved more than what classical cryptography itself can provide.

A new classical bit commitment protocol taking advantage of cryptographic constraints imposed by special relativity was proposed by Kent.[18] The author shows that the protocol is unconditionally secure against both classical or quantum attacks. Subsequently, Kent proposed an unconditionally secure QBC protocol that uses quantum and relativistic effects and that requires 4 additional trusted agents.[19] This protocol was recently implemented by two different groups. In one implementation the thrusted agents were separated by more than 9000 km, corresponding to a commitment time of 15.6 ms,[8] and in the second implementation the agents were separated by about 20 km, achieving a commitment time of about 60 μs and a cheating probability less than 5.68×10^{-2}.[3]

Although not unconditionally secure, an experimental demonstration of a practical QBC protocol whose security is based on current technological limitations was presented by Danan and Vaidman.[20] The technological limitations in question are the lack of non-demolition measurements and long-term stable quantum memories. The protocol is based on BB84 scheme and has the advantages of being relatively simple to implement and of immediate realization using current technology. Quantum non-demolition (QND) measurements are under study basically since the foundation of quantum theory.[21] If Alice was able to perform a non-demolition measurement, she would be able to detect the presence of a photon without destroying it or affecting its polarization state. However, practical QND measurements are still in its preliminary stage. An experimental demonstration of 90% QND measurements in a controlled environment (photons confined in a cavity) was presented, although this type of measurements is more suitable for computational purposes, rather than cryptographic.[22] Even if Alice

had access to QND measurements, she would need a stable long-term quantum memory in order to perform the measurements later in time. Advances in quantum-memories technology are also significant and we step from a 2 seconds proposal[23] to a 39 minutes one[24] only in two years. Although notable, current technology allows only noisy memories, usually implemented in optical fibers, increasing the probability to have errors in the system, since the qubit losses increase exponentially with the length of the fiber.

Recently, a two-state version of the protocol presented by Danan and Vaidman[20] was proposed.[25] This paper addresses several issues related to a practical implementation, e.g. modeling of errors due to noise and equipment imperfections, a quantitative analysis of the effects of noise in the system, the optimal cheating strategy or the protocol's security. An experimental implementation based on this work[25] was recently demonstrated.[26] The security of the implementation of this protocol[26] is based on the lack of long-term stable quantum memories.

2.2 Applications

Bit commitment is a building block for several cryptographic primitives, including coin tossing,[13,27,28] zero-knowledge proofs,[29,30] oblivious transfer[31,32] and secure two-party computation.[33] In terms of real-world applications it will be useful in a near future. High-speed trading stock market is one example, secure voting or long-distance gambling are other applications.[8,34]

Since the discovery of Mayers in 1997, efforts on bit commitment were put in combining quantum theory with relativity and in theoretical analysis and experimental realizations of practical protocols with noisy/bounded memories.[6,32,35–38] It was shown that if the memory is not ideal but has a finite noise, there are protocols whose security will not be compromised as long as the time difference between the commitment and the opening phase is bigger than some threshold value, determined by the noisy characteristics of the memory used.[25] Regarding protocols based on relativistic effects, Kent has been working in this topic for several years, proposing several protocols for unconditionally secure bit commitment.[7,18,19,39,40]

Nowadays, with the recent experimental demonstrations of unconditionally secure bit commitment,[3,8] a broad range of opportunities is now open for a lot of applications.

3. EXPERIMENTAL QUANTUM BIT COMMITMENT USING TECHNOLOGICAL LIMITATIONS

In Loura et. al,[25] the authors have proposed a nonrelativistic two-state QBC protocol whose security relies on current technological limitations. In fact there are some disadvantages of using relativistic bit commitment protocols that makes it worthwhile to continue to study nonrelativistic protocols. First, contrary to relativistic bit commitment, the protocol in Loura et. al[25] uses less equipment and is less costly, making it easier to implement. Second, in relativistic protocols both parties need to be in separated and secure locations and the communication between them needs to be continuous during the entire commitment phase.[41] In particular in the protocol proposed by Kent, 4 thrusted and secure agents are needed, which increases the complexity and cost of the system.[19]

Next we present an implementation of the protocol presented in Loura et. al.[25] The protocol has three phases, initialization, commitment and revealing, which run as follows:

1. **Initialization:** Bob generates a random sequence of classical bits, encodes them in one of two states of polarization, $|0\rangle$ or $|1\rangle$ and sends them to Alice. The two states of polarization are not orthogonal, i.e. $\langle 0|1\rangle = \cos(\pi/4)$.

2. **Commitment:** Alice performs the measurements right after receiving a photon from Bob, choosing one of two orthogonal observables, that we called $\hat{C}_0$ and $\hat{C}_1$, on all photons.

3. **Revealing:** Alice reveals her commitment by informing Bob about the observable she measured and the results obtained.

One can define conditional probabilities for the optical contributions ($p_c(r|b)$, with $c, r, b \in \{0, 1\}$) that a result r is obtained when measuring the observable $\hat{C}_c$ on state $|b\rangle$.[25]

- If Alice measures $\hat{C}_0$:

$$\begin{aligned} p_0(0|0) &= 1 - \frac{p}{2} \\ p_0(1|0) &= \frac{p}{2} \\ p_0(0|1) &= 1/2 \\ p_0(1|1) &= 1/2\,. \end{aligned} \tag{1}$$

- If Alice measures $\hat{C}_1$:

$$\begin{aligned} p_1(0|0) &= 1/2 \\ p_1(1|0) &= 1/2 \\ p_1(0|1) &= \frac{p}{2} \\ p_1(1|1) &= 1 - \frac{p}{2}\,. \end{aligned} \tag{2}$$

The parameter p in (1) and (2) traduces the amount of optical noise in a depolarizing channel model of white noise.

In Fig. 1 we present the experimental setup used to implement the two-state QBC described above. On Bob's

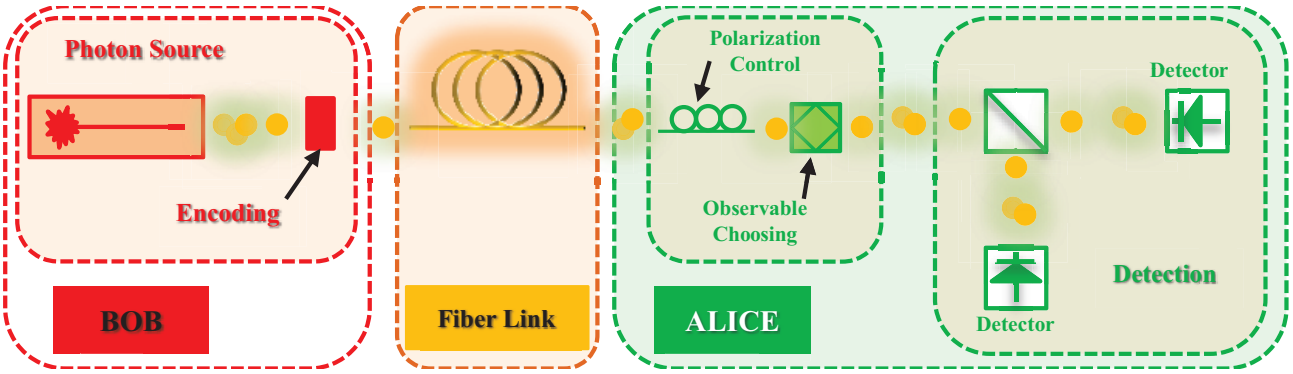


Figure 1. Experimental scheme to implement the two-state QBC protocol.

side, a laser at 1550 nm band was attenuated to generate an average number of photons per pulse much lower than 1. Next, photons were encoded into polarization and sent through a 1 km long fiber. At the receiver's side, Alice used a polarization controller to compensate for random rotations of polarization inside the optical fiber and performed her commitment. The commitment to an observable was done using a wave plate. Photons were detected using two avalanche photodiodes. The experimental results of the measurement probability obtained in five different runs when measuring $\hat{C}_0$ and $\hat{C}_1$ are shown in Fig. 2. The average measurement probabilities obtained from each observable are presented in Table 1. Note that measurements in equal bases means that

Table 1. Average measurement probabilities from $\hat{C}_0$ and $\hat{C}_1$.

	Equal bases (%)	Different bases (%)
$\hat{C}_0$	97.69	50.10
$\hat{C}_1$	97.55	51.83

the measurement observable coincides with the basis from which the state was prepared and measurements in different bases means that the two are different.

The results obtained are according to theory in (1) and (2), since in the ideal case (where there is no optical noise, $p = 0$), when the measurements are performed in equal bases the success rates (given by $p_0(1|1)$ and

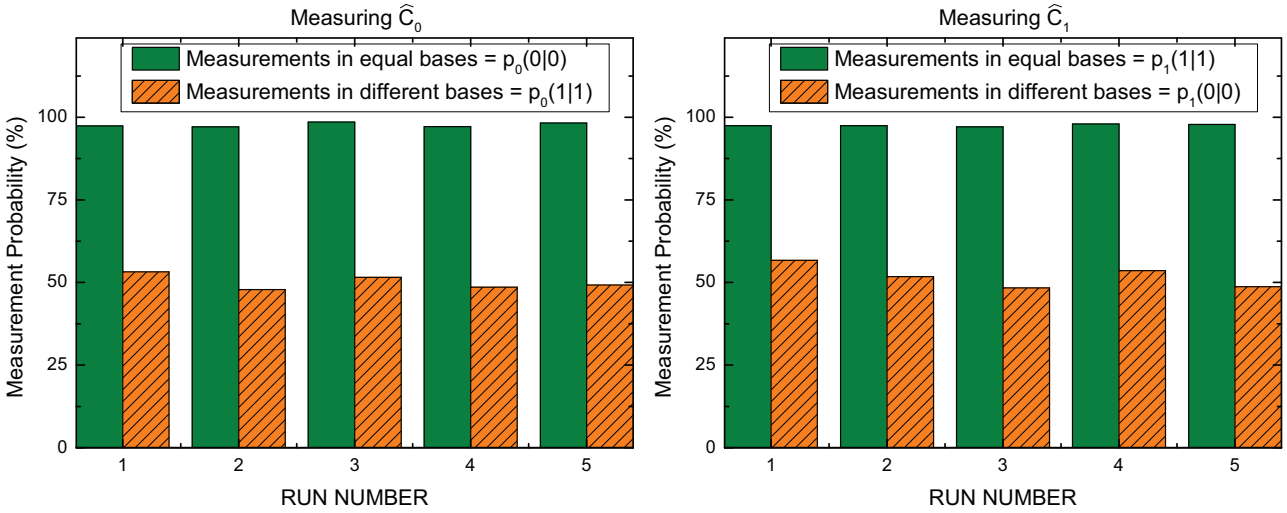


Figure 2. Experimental results for the measurement probability in $\hat{C}_0$ and $\hat{C}_1$, obtained for five different runs.

$p_1(0|0)$) should be 100%, while for measurements in different bases (given by $p_0(0|0)$ and $p_1(1|1)$) the rates should be 50%. Since the alignment between Bob and Alice is not perfect, we will get wrong detections and consequently a rate lower than 100%, or different from 50%, depending if we are talking about measurements in equal bases or different bases, respectively.

The wrong counts when we are measuring in equal bases will contribute to the quantum-bit error rate (QBER) of the system, which is one of the most important parameters to quantify the quality of the transmission. The overall QBER, including optical and non-optical contributions can be written as,[42]

$$\mathrm{QBER} = \frac{\mu t_{\mathrm{link}} \eta P_{\mathrm{opt}} + P_{\mathrm{dark}}}{\mu t_{\mathrm{link}} \eta + 2P_{\mathrm{dark}}}\,. \tag{3}$$

The optical contribution to errors due to rotation of polarization is given by,

$$P_{\mathrm{opt}} = p_0(1|0) = p_1(0|1) = \frac{1-V}{2}, \tag{4}$$

where V is the visibility of the state of polarization. For the non-optical part, the channel contribution from photon absorption is,

$$t_{\mathrm{link}} = 10^{\frac{-\alpha L}{10}}, \tag{5}$$

where α is the fiber attenuation coefficient [in dB/km] and L is the fiber length [in km]. The parameter μ is the average number of photons per pulse and P_{dark} and η are the detectors' dark counts and quantum efficiency, respectively. The success rate (SRATE) in the measurement, i.e., when the measurement basis is equal to the coding basis, can be defined as,

$$\mathrm{SRATE} = 1 - \mathrm{QBER}. \tag{6}$$

In Fig. 3 we show the success rate as a function of the fiber length obtained from (6). From this figure it can be seen that, according to the parameters used and which are presented in Table 2, bit commitment can be securely performed even if the parties are separated through a quantum channel of about 175 km. This is due to the fact that the success rates for the optimal cheating strategy is $\cos^2(\pi/8) = 0.8536$.[25] Thus, if the errors induced by an honest part (Alice) are identical to the error of a cheating one, the two strategies are indistinguishable and the protocol cannot be performed.

4. CONCLUSIONS

We reviewed the state of the art of bit commitment protocol. It was verified that bit commitment is a very important primitive in cryptography and has seen a rapid progress. Nowadays, there are proposals for unconditionally secure bit commitment which are based in special relativity combined with quantum mechanics. Some

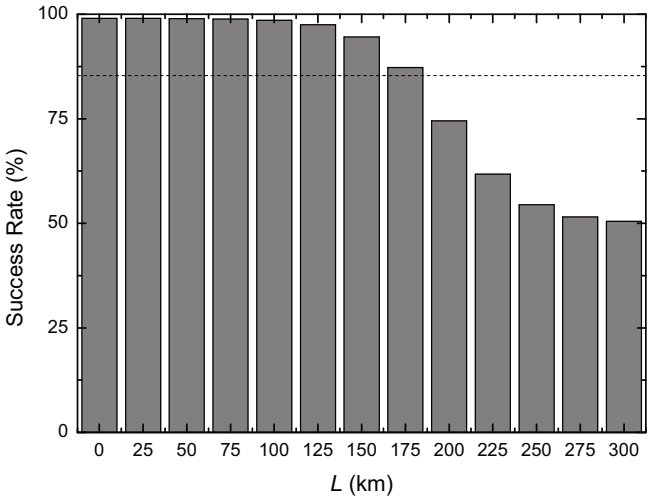


Figure 3. Success rate as a function of the fiber length. The dashed line sets the threshold to secure bit commitment.

Table 2. Parameters used in (6) to plot Fig. 3.

Parameter	Value
μ	0.2
α	0.2 dB/km
η	0.1
P_{dark}	1×10^{-6}
V	0.98

of these proposals were already implemented experimentally. Regarding the one presented in this work, which is of relatively simplicity of implementation if compared with relativistic options, it was demonstrated that it can be achieved secure QBC. The measurement probabilities when measuring each one of the observables were far above the minimum theoretical security limit of 85.36%. Also, it was shown that it is worthwhile to continue to study non-relativistic QBC protocols while relativistic ones still present some disadvantages. Bit commitment has a lot of applications and can be a serious candidate to a real-world implementation in a near future.

ACKNOWLEDGMENTS

This work was supported in part by the FCT - Fundação para a Ciência e a Tecnologia, through the PhD Grants SFRH/BD/79482/2011 and SFRH/BD/79571/2011; the Post Doctoral Grants SFRH/BPD/77286/2011 and SFRH/BPD/96457/2013; by the Conselho de Reitores das Universidades Portuguesas (CRUP) project 'Ação Integrada E 91/12' and by the FCT and the Instituto de Telecomunicações under the PEst-OE/EEI/LA0008/2013 program, project "CV-Quantum".

REFERENCES

[1] Crandall, R. and Pomerance, C., [*Prime Numbers: A Computational Perspective*], Springer, 2nd ed. (Aug. 2005).

[2] Vaudenay, S., [*A Classical Introduction to Cryptography: Applications for Communications Security*], Springer (2006).

[3] Liu, Y., Cao, Y., Curty, M., Liao, S.-K., Wang, J., Cui, K., Li, Y.-H., Lin, Z.-H., Sun, Q.-C., Li, D.-D., Zhang, H.-F., Zhao, Y., Chen, T.-Y., Peng, C.-Z., Zhang, Q., Cabello, A., and Pan, J.-W., "Experimental unconditionally secure bit commitment," *Phys. Rev. Lett.* **112**, 010504 (Jan. 2014).

[4] Mayers, D., "Unconditionally Secure Quantum Bit Commitment is Impossible," *Phys. Rev. Lett.* **78**, 3414–3417 (Apr. 1997).

[5] Lo, H.-K. and Chau, H. F., "Is Quantum Bit Commitment Really Possible?," *Phys. Rev. Lett.* **78**, 3410–3413 (Apr. 1997).

[6] Wehner, S., Curty, M., Schaffner, C., and Lo, H.-K., "Implementation of two-party protocols in the noisy-storage model," *Phys. Rev. A* **81**, 052336 (May 2010).

[7] Kent, A., "Unconditionally secure bit commitment with flying qudits," *New J. Phys.* **13**, 113015 (Nov. 2011).

[8] Lunghi, T., Kaniewski, J., Bussières, F., Houlmann, R., Tomamichel, M., Kent, A., Gisin, N., Wehner, S., and Zbinden, H., "Experimental bit commitment based on quantum communication and special relativity," *Phys. Rev. Lett.* **111**, 180504 (Nov. 2013).

[9] Shannon, C., "A mathematical theory of communication," *Bell System Technical Journal* **27**, 379–423, 623–656 (1948).

[10] Data Encryption Standard (DES), "Federal Information Processing Standards Publication 46-3. U.S. Department of Commerce, National Institute of Standards and Technology," (1999).

[11] Advanced Encryption Standard (AES), "Federal Information Processing Standards Publication 197. U.S. Department of Commerce, National Institute of Standards and Technology," (2001).

[12] Wiesner, S., "Conjugate coding," *SIGACT News* **15**(1), 78–88 (1983).

[13] Blum, M., "Coin flipping by telephone.," in [*CRYPTO*], Gersho, A., ed., 11–15, U. C. Santa Barbara, Dept. of Elec. and Computer Eng., ECE Report No 82-04 (1981).

[14] Brassard, G., Chaum, D., and Crépeau, C., "Minimum disclosure proofs of knowledge.," *J. Comput. Syst. Sci.* **37**(2), 156–189 (1988).

[15] Bennett, C. H. and Brassad, G., "Quantum cryptography: Public-key distribution and coin tossing," *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* , 175–179 (Dec. 1984).

[16] Brassard, G., Crépeau, C., Jozsa, R., and Langlois, D., "A quantum bit commitment scheme provably unbreakable by both parties," in [*Foundations of Computer Science, 1993. Proceedings., 34th Annual Symposium on*], 362–371 (1993).

[17] Crépeau, C., [*What is going on with Quantum Bit Commitment?*], Czech Technical University Publishing House (1996).

[18] Kent, A., "Unconditionally secure bit commitment," *Phys. Rev. Lett.* **83**, 1447–1450 (Aug. 1999).

[19] Kent, A., "Unconditionally secure bit commitment by transmitting measurement outcomes," *Phys. Rev. Lett.* **109**, 130501 (Sept. 2012).

[20] Danan, A. and Vaidman, L., "Practical Quantum Bit Commitment Protocol," *Quantum Information Processing* **11**, 769–775 (June 2012).

[21] Braginsky, V. B. and Khalili, F. Y., "Quantum nondemolition measurements: the route from toys to tools," *Rev. Mod. Phys.* **68**, 1–11 (Jan. 1996).

[22] Johnson, B. R., Reed, M. D., Houck, A. A., Schuster, D. I., Bishop, L. S., Ginossar, E., Gambetta, J. M., Dicarlo, L., Frunzio, L., Girvin, S. M., and Schoelkopf, R. J., "Quantum non-demolition detection of single microwave photons in a circuit," *Nature Physics* **6**, 663–667 (Sept. 2010).

[23] Maurer, P. C., Kucsko, G., Latta, C., Jiang, L., Yao, N. Y., Bennett, S. D., Pastawski, F., Hunger, D., Chisholm, N., Markham, M., Twitchen, D. J., Cirac, J. I., and Lukin, M. D., "Room-Temperature Quantum Bit Memory Exceeding One Second," *Science* **336**, 1283–1286 (June 2012).

[24] Saeedi, K., Simmons, S., Salvail, J. Z., Dluhy, P., Riemann, H., Abrosimov, N. V., Becker, P., Pohl, H.-J., Morton, J. J. L., and Thewalt, M. L. W., "Room-temperature quantum bit storage exceeding 39 minutes using ionized donors in silicon-28," *Science* **342**(6160), 830–833 (2013).

[25] Loura, R., Almeida, A. J., André, P. S., Pinto, A. N., Mateus, P., and Paunkovic, N., "Noise and measurement errors in a practical two-state quantum bit commitment protocol," *Phys. Rev. A* (2014).

[26] Almeida, A. J., Stojanovic, A. D., Paunković, N., Loura, R., Muga, N. J., Silva, N. A., Mateus, P., André, P. S., and Pinto, A. N., "Implementation of a Two-State Quantum Bit Commitment Protocol in Optical Fibers," *submitted to Quantum Information Processing* (2014).

[27] Damgård, I. and Lunemann, C., "Quantum-secure coin-flipping and applications," in [*Advances in Cryptology ASIACRYPT 2009*], Matsui, M., ed., *Lecture Notes in Computer Science* **5912**, 52–69, Springer Berlin Heidelberg (2009).

[28] Molina-Terriza, G., Vaziri, A., Ursin, R., and Zeilinger, A., "Experimental quantum coin tossing," *Phys. Rev. Lett.* **94**, 040501 (Jan. 2005).

[29] Brassard, G., "Modern cryptology: A tutorial," (1998).

[30] Kobayashi, H., "General properties of quantum zero-knowledge proofs," in [*Theory of Cryptography*], Canetti, R., ed., *Lecture Notes in Computer Science* **4948**, 107–124, Springer Berlin Heidelberg (2008).

[31] Li, Y.-B., Wen, Q.-Y., Qin, S.-J., Guo, F.-Z., and Sun, Y., "Practical quantum all-or-nothing oblivious transfer protocol," *Quantum Information Processing* **13**, 131–139 (Jan. 2014).

[32] Erven, C., Ng, N., Gigov, N., Laflamme, R., Wehner, S., and Weihs, G., "An experimental implementation of oblivious transfer in the noisy storage model," *Nature Communications* **5** (Mar. 2014).

[33] Lindell, Y. and Pinkas, B., "An efficient protocol for secure two-party computation in the presence of malicious adversaries," in [*Advances in Cryptology - EUROCRYPT 2007*], Naor, M., ed., *Lecture Notes in Computer Science* **4515**, 52–78, Springer Berlin Heidelberg (2007).

[34] Broadbent, A. and Tapp, A., "Information-theoretically secure voting without an honest majority." Cryptology ePrint Archive, Report 2008/266 (2008).

[35] Wehner, S., Schaffner, C., and Terhal, B. M., "Cryptography from noisy storage," *Phys. Rev. Lett.* **100**, 220502 (June 2008).

[36] Schaffner, C., Terhal, B., and Wehner, S., "Robust cryptography in the noisy-quantum-storage model," *Quantum Info. Comput.* **9**, 963–996 (Nov. 2009).

[37] Konig, R., Wehner, S., and Wullschleger, J., "Unconditional security from noisy quantum storage," *Information Theory, IEEE Transactions on* **58**, 1962–1984 (Mar. 2012).

[38] Ng, N. H. Y., Joshi, S. K., Chen Ming, C., Kurtsiefer, C., and Wehner, S., "Experimental implementation of bit commitment in the noisy-storage model," *Nature Communications* **3** (Dec. 2012).

[39] Hardy, L. and Kent, A., "Cheat sensitive quantum bit commitment," *Phys. Rev. Lett.* **92**, 157901 (Apr. 2004).

[40] Kent, A., "Quantum tasks in Minkowski space," *Classical and Quantum Gravity* **29**, 224013 (Nov. 2012).

[41] Kent, A., "Quantum Bit String Commitment," *Phys. Rev. Lett.* **90**, 237901 (June 2003).

[42] Kumavor, P. D., Beal, A. C., Yelin, S., Donkor, E., and Wang, B. C., "Comparison of Four Multi-User Quantum Key Distribution Schemes Over Passive Optical Networks," *Journal of Lightwave Technology* **23**, 268 (Jan. 2005).