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

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.

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

$$p_0(1|0) = p_1(0|1) = p/2. \qquad (7)$$

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

$$\mathcal{E}_d(\hat{\rho}) = (1-p)\hat{\rho} + p\frac{\hat{I}}{2} = \left(1 - \frac{3}{4}\right)\hat{\rho} + \frac{p}{4}\sum_{i=1}^{3}\hat{\sigma}_i\hat{\rho}\hat{\sigma}_i, \qquad (8)$$

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 (η = QBER represents the optical part of the quantum-bit error rate):

(i) Alice measures $\hat{C}_0$:

$$p_0(0|0) = 1 - \frac{p}{2} = 1 - \eta, \quad p_0(1|0) = \frac{p}{2} = \eta,$$

$$p_0(0|1) = (1-p)\cos^2\theta + \frac{p}{2} = (1-2\eta)\cos^2\theta + \eta,$$

$$p_0(1|1) = (1-p)\sin^2\theta + \frac{p}{2} = (1-2\eta)\sin^2\theta + \eta. \qquad (9)$$

[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 $n \leqslant N$).

(ii) Alice measures $\hat{C}_1$:

$$\begin{aligned}
p_1(0|0) &= (1-p)\sin^2\theta + \frac{p}{2} = (1-2\eta)\sin^2\theta + \eta,\\
p_1(1|0) &= (1-p)\cos^2\theta + \frac{p}{2} = (1-2\eta)\cos^2\theta + \eta,\\
p_1(0|1) &= \frac{p}{2} = \eta, \quad p_1(1|1) = 1 - \frac{p}{2} = 1-\eta. \qquad (10)
\end{aligned}$$

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

$$\mathcal{E}_{b_0}(\hat{\rho}) = (1-p)\hat{\rho} + p\hat{\sigma}_{x_0}\hat{\rho}\hat{\sigma}_{x_0}. \qquad (11)$$

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 $\sigma_x = [\hat{\sigma}_{x_0}]_{\mathcal{B}_0} = [\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}]$. 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 $\sigma_z = [\hat{\sigma}_{z_0}]_{\mathcal{B}_0} = [\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}]$. The operator-sum representation of the phase-flip channel is

$$\mathcal{E}_{p_0}(\hat{\rho}) = (1-p)\hat{\rho} + p\hat{\sigma}_{z_0}\hat{\rho}\hat{\sigma}_{z_0}. \qquad (12)$$

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

$$\mathcal{E}_u(\hat{\rho}) = \hat{U}\hat{\rho}\hat{U}^\dagger, \qquad (13)$$

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),

$$[\hat{U}]_{\mathcal{B}_0} = \begin{bmatrix} e^{i\lambda}\cos\alpha & -e^{-i\mu}\sin\alpha \\ e^{i\mu}\sin\alpha & e^{-i\lambda}\cos\alpha \end{bmatrix}, \qquad (14)$$

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

$$
\begin{aligned}
\langle F(|0\rangle,|1\rangle)\rangle = 1/2[&F(p_0(*|0),p_0(*|1))\\
&+F(p_1(*|0),p_1(*|1))].
\end{aligned}
\tag{19}
$$

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

$$
\begin{aligned}
\langle F(\hat{C}_0,\hat{C}_1)\rangle = 1/2[&F(p_0(*|0),p_1(*|0))\\
&+F(p_0(*|1),p_1(*|1))],
\end{aligned}
\tag{20}
$$

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

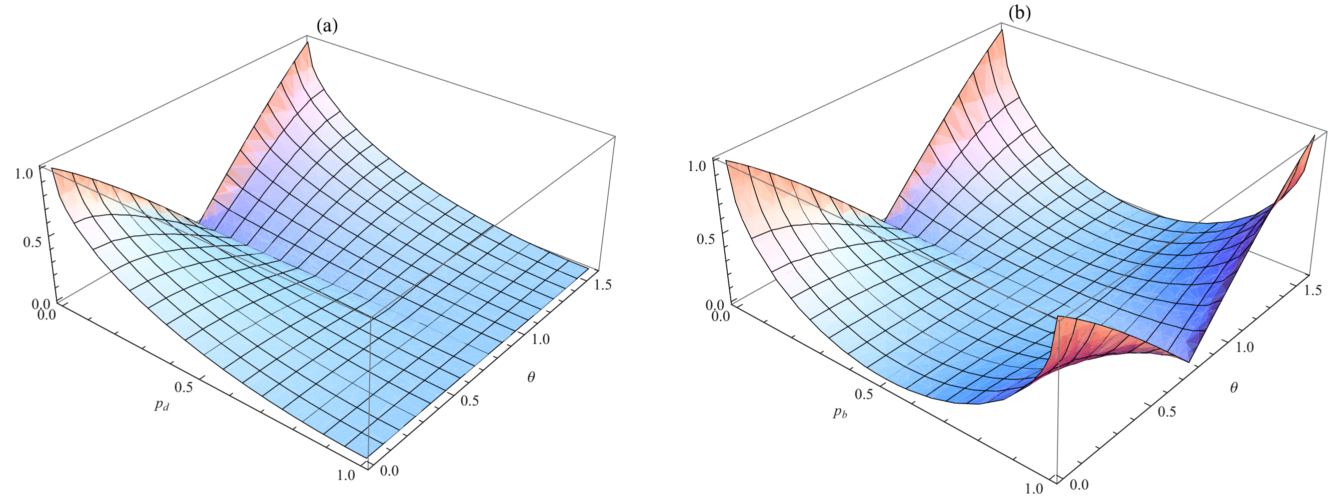

FIG. 1. (Color online) (a) Graph of $|\langle F(|0\rangle,|1\rangle)\rangle - \langle F(\hat{C}_0,\hat{C}_1)\rangle|$ as a function of θ and p_d. (b) Graph of $|\langle F(|0\rangle,|1\rangle)\rangle - \langle F(\hat{C}_0,\hat{C}_1)\rangle|$ as a function of θ and p_b.

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.

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

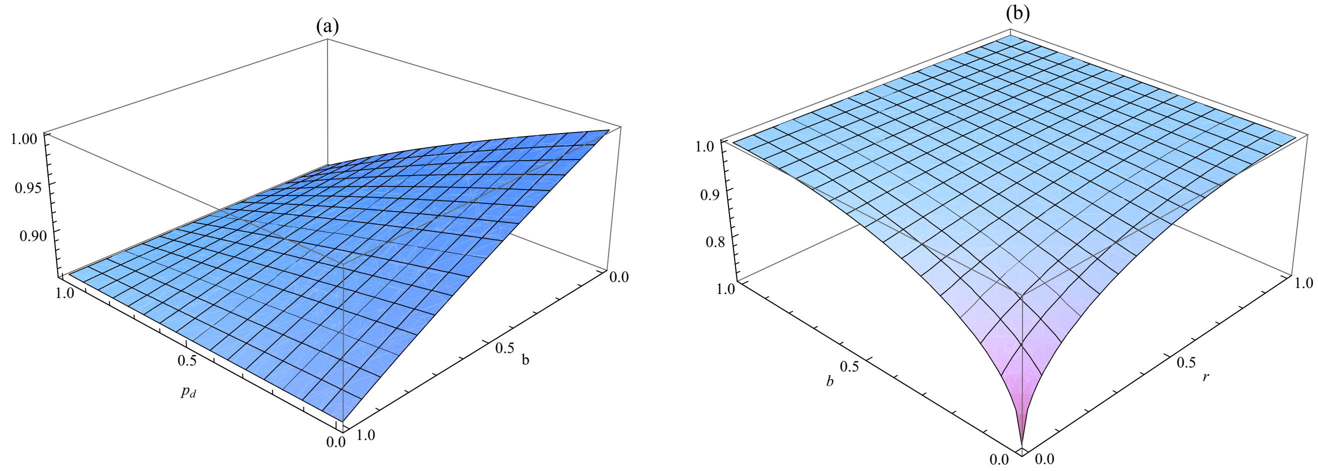

FIG. 2. (Color online) Fidelity $\langle F(\mathcal{E})\rangle$ between noisy and noiseless channels (a), and fidelity $\langle F(|0\rangle,|1\rangle)\rangle = \langle F(\hat{C}_0,\hat{C}_1)\rangle$ within the noisy channel (b), as functions of the depolarizing coefficient p_d and the joint (preparation- and measurement-induced) bit-flip coefficient $b = 1-(1-2p_b^p)(1-2p_b^m)$.

[6] Indeed, its infinitesimal form is a true metric, the so-called Fisher information metric (see, for example, [34], p. 87).

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

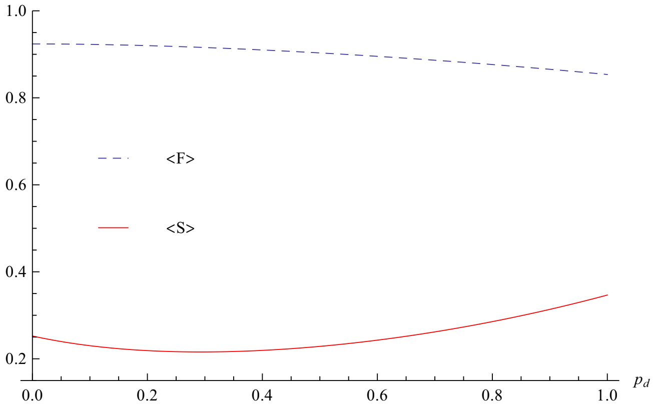

FIG. 3. (Color online) Fidelity $\langle F(\mathcal{E})\rangle$ [dashed (blue) line] and relative entropy $\langle S(\mathcal{E})\rangle$ [full (red) line] between an honest strategy without the presence of noise, and the optimal cheating strategy in the presence of white noise, as a function of the 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

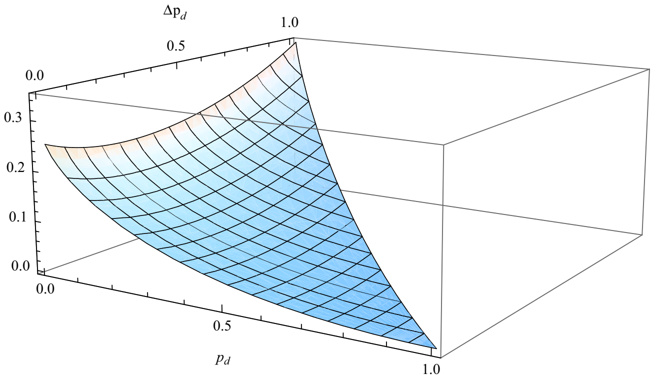

FIG. 4. (Color online) Introducing a small noise Δp_d helps a dishonest party.

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 $(p_0(*|0), p_{\mathrm{ch}}^{\mathcal{E}}(*|0))$ and $(p_1(*|1), p_{\mathrm{ch}}^{\mathcal{E}}(*|1))$ are, but also in the distinguishability of the cross terms given by the pairs $(p_0(*|1), p_{\mathrm{ch}}^{\mathcal{E}}(*|1))$ and $(p_1(*|0), p_{\mathrm{ch}}^{\mathcal{E}}(*|1))$ 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 $p_{\mathrm{ch}}^{\mathcal{E}}(*|0)$ and $p_{\mathrm{ch}}^{\mathcal{E}}(*|1)$ 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):

$$P(q||p_0) > \alpha \Rightarrow P(q||p_{\mathrm{ch}}) < \beta. \tag{28}$$

Note that it is more likely to produce statistics that look as if they were obtained by committing to 0 by measuring $\hat{C}_{\mathrm{ch}}$ 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 $\hat{C}_{\mathrm{ch}}$. 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

$$P(q(*|0)||p_{\mathrm{ch}}(*|0)) < \beta \quad \text{and} \quad P(q(*|1)||p_{\mathrm{ch}}(*|1)) < \beta'. \tag{30}$$

(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 μ_{ch} and σ_{ch}, and we can define β to be $\beta_{2\sigma} = \Phi(\mu_{\mathrm{ch}} - 2\sigma_{\mathrm{ch}}; \mu_{\mathrm{ch}}, \sigma_{\mathrm{ch}})$, for which the statistics obtained by measuring $\hat{C}_{\mathrm{ch}}$ 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 $P(q(*|1)||p_{\mathrm{ch}}(*|1))$ (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., $\binom{n}{n_0}$.

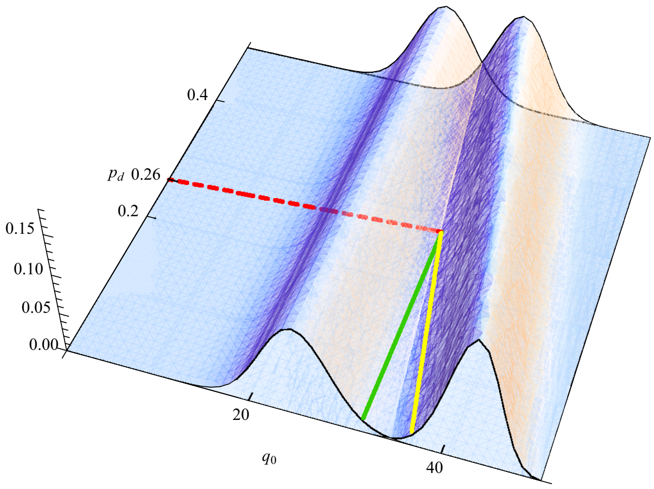

FIG. 5. (Color online) $P(q(*|1)||p_0(*|1))$ (left bump) and $P(q(*|1)||p_{\rm ch}(*|1))$ (right bump), as functions of a depolarizing coefficient p_d and $q_0 = \frac{n(0|1)}{n(0)}$ for $n(0) = 50$. The green (left) full line represents the intersection between the left bump and the $q_0 = \mu_0 + 2\sigma_0$ 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)

$$\begin{aligned} &P(q(*|0)||p_0^{\Delta t}(*|0)) < \beta^{\Delta t} \quad \text{and} \\ &P(q(*|1)||p_0^{\Delta t}(*|1)) < \beta^{\Delta t}. \end{aligned} \tag{34}$$

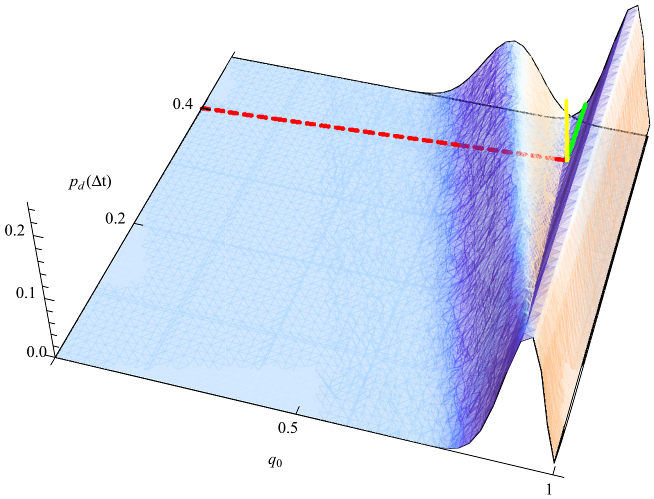

FIG. 6. (Color online) $P(q(*|0)||p_0(*|0))$ (right bump) and $P(q(*|0)||p_0^{\Delta t}(*|0))$ (left bump) as a function of $p_d(\Delta t)$ and $q_0 = \frac{n(0|1)}{n(0)}$ for a fixed value of $p_d = 0.15$ and $n(0) = 50$. The green (right) full line represents the intersection between the right bump 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, $\text{QBER} = \delta p_0(1|0) = \delta p_1(0|1)$, 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 $N_{\rm tot}$ be the total number of qubits received in the state $|0\rangle$, and let $N_{\rm wrong}$ 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

$$\mathrm{QBER} = \frac{N_{\rm wrong}}{N_{\rm tot}} = \frac{\frac{N_{\rm wrong}}{T}}{\frac{N_{\rm tot}}{T}} = \frac{R_{\rm error}}{R_{\rm tot}}. \qquad \text{(B1)}$$

Here $R_{\rm tot} = N_{\rm tot}/T$ and $R_{\rm error} = N_{\rm wrong}/T$ are the total and the error rates, respectively, for the qubits received in the state $|0\rangle$. If $R_{\rm raw}$ is the overall source rate, including both $|0\rangle$ and $|1\rangle$ states, then

$$R_{\rm tot} = \tfrac{1}{2} R_{\rm raw}, \qquad \text{(B2)}$$

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 $f_{\rm rep}\mu$) that were not absorbed and that managed to arrive to detectors (given by $t_{\rm link}$) and were detected (given by η) is

$$R_{\rm raw} = f_{\rm rep}\mu t_{\rm link}\eta\,. \qquad \text{(B3)}$$

Here, $f_{\rm rep}$ 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, $f_{\rm rep}\mu$ is the number of photons sent, in the unit of time. The probability that a sent photon arrives at a detector is $t_{\rm link} \sim 10^{-\alpha L}$ (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, $R_{\rm error} = R_{\rm opt} + R_{\rm det}$, but here we are only interested in the error arising due to dark counts. We have

$$R_{\rm det} = \tfrac{1}{2}\left(\tfrac{1}{2} f_{\rm rep}\right)(1 - \mu t_{\rm link}\eta)p_{\rm dark} \approx \tfrac{1}{4} f_{\rm rep} p_{\rm dark}\,. \qquad \text{(B4)}$$

Here, $\frac{1}{2}$ is the probability that a wrong detector will click, $(\frac{1}{2} f_{\rm rep})$ is the rate of photons sent in the "right" state (in our case $|0\rangle$), and $(1 - \mu t_{\rm link}\eta)$ 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, $\mu t_{\rm link}\eta \ll 1$. Also, we neglect the less probable cases of "right" dark counts, when a photon does not arrive at detectors.

Thus, we get

$$\begin{aligned} \delta p_0(1|0) = \delta p_1(0|1) &= \frac{p_{\rm dark}}{2\mu t_{\rm link}\eta}, \\ \delta p_0(0|0) = \delta p_1(1|1) &= -\frac{p_{\rm dark}}{2\mu t_{\rm link}\eta}. \end{aligned} \qquad \text{(B5)}$$

[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).