

## Fair loss-tolerant quantum coin flipping

Guido Berlín,<sup>1,\*</sup> Gilles Brassard,<sup>1,†</sup> Félix Bussi  res,<sup>2,3,‡</sup> and Nicolas Godbout<sup>2,§</sup><sup>1</sup>*D  partement d'informatique et de recherche op  rationnelle, Universit   de Montr  al, CP 6128, Succursale Centre-Ville, Montr  al, Qu  bec, Canada H3C 3J7*<sup>2</sup>*Laboratoire des fibres optiques, D  partement de g  nie physique,   cole Polytechnique de Montr  al, CP 6079, Succursale Centre-Ville, Montr  al, Qu  bec, Canada H3C 3A7*<sup>3</sup>*Department of Physics and Astronomy and Institute for Quantum Information Science, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4*

(Received 4 May 2009; published 10 December 2009)

Coin flipping is a cryptographic primitive in which two spatially separated players, who do not trust each other, wish to establish a common random bit. If we limit ourselves to classical communication, this task requires either assumptions on the computational power of the players or it requires them to send messages to each other with sufficient simultaneity to force their complete independence. Without such assumptions, all classical protocols are so that one dishonest player has complete control over the outcome. If we use quantum communication, on the other hand, protocols have been introduced that limit the maximal bias that dishonest players can produce. However, those protocols would be very difficult to implement in practice because they are susceptible to realistic losses on the quantum channel between the players or in their quantum memory and measurement apparatus. In this paper, we introduce a quantum protocol and we prove that it is completely impervious to loss. The protocol is *fair* in the sense that either player has the same probability of success in cheating attempts at biasing the outcome of the coin flip. We also give explicit and optimal cheating strategies for both players.

DOI: [10.1103/PhysRevA.80.062321](https://doi.org/10.1103/PhysRevA.80.062321)

PACS number(s): 03.67.Dd, 03.67.Hk, 89.70.  a

## I. INTRODUCTION

Coin flipping by telephone was first introduced with these words by Blum in 1981: "Alice and Bob [...] have just divorced, live in different cities, want to decide who gets the car" [1]. They agree that the best thing to do is to flip a coin, but neither of them trusts the other and they are unable to agree on a mutually trusted third party to do the flip for them. More generally, coin-flipping protocols (also known as "coin tossing") can be used whenever two players need to pick a random bit even though it could be to the advantage of one of them (or perhaps both) to choose, or at least bias, the outcome of the protocol.

The original coin-flipping protocol introduced by Blum is *asynchronous* in the sense that it consists of a sequence of rounds in which the two players alternate in sending messages to each other. The security of Blum's protocol depends on the assumed difficulty of factoring large numbers. Such an assumption is of course of little value in our quantum world, owing to Shor's algorithm [2], but classical coin flipping can be based on more general one-way functions, which could potentially be immune to quantum attacks. Nevertheless, any coin-flipping protocol that takes place by the asynchronous transmission of classical messages has the property that one of the players has complete control over the outcome, given sufficient computing power. In the best case, such protocols can be computationally secure and even *that*

depends on unproven computational complexity assumptions.

Unconditionally secure classical coin-flipping protocols are possible in the *synchronous* model, in which the players are requested to send messages to each other with sufficient simultaneity to force their complete independence. Such protocols are called *relativistic* because special relativity must be invoked to prevent Alice from waiting to receive Bob's message before choosing her own (and vice versa). Relativistic protocols must be implemented carefully because their security depends crucially on the physical distance between the players, and either of them could try to fool the other by pretending to be farther away than they really are. Such cheating attempts can be thwarted if each player has a trusted agent near the other player [3]. For the rest of this paper, we only consider asynchronous protocols and "coin-flipping protocol" will systematically mean "asynchronous coin-flipping protocol."

In *quantum* coin-flipping protocols, Alice and Bob are allowed to exchange quantum states. Such protocols were first investigated in 1984 by Bennett and Brassard [4]. In that paper, a protocol was presented and it was shown that "ironically (it) can be subverted by use of a still subtler quantum phenomenon, the Einstein-Podolsky-Rosen paradox," making it the first use of entanglement [5] in quantum cryptography. We shall refer to it henceforth as the "BB84 protocol" (Bennett-Brassard 1984 protocol), not to be confused with the better-known quantum key distribution protocol introduced in the same paper. The question was left open: can there be a perfect quantum coin-flipping protocol? The proof that this is impossible was given more than a decade later by Lo and Chau [6], whose result was further clarified by Mayers *et al.* [7]. Nevertheless, if quantum coin-flipping proto-

\*berlingu@iro.umontreal.ca

†brassard@iro.umontreal.ca

‡felix.bussi  res@polymtl.ca

§nicolas.godbout@polymtl.ca

cols cannot be perfect, can they at least be better than anything classically possible?

To make this question more precise, we say that one player enjoys an  $\varepsilon$  bias if a cheating strategy exists by which that player could choose either bit and influence the outcome of the protocol to be that bit with probability at least  $\frac{1}{2} + \varepsilon$ , assuming that the other player follows the protocol honestly. This definition is *unconditional* in the sense that we allow the would-be cheater to enjoy unlimited computational power and a technology limited only by the known laws of physics. The *bias* of a protocol is the largest value of  $\varepsilon$  so that at least one player enjoys an  $\varepsilon$  bias. A *perfect* protocol would be one whose bias is 0, but they cannot exist, classical or quantum. At the other end of the spectrum, a protocol whose bias is 0.5 is considered to be *completely broken*. All classical protocols are completely broken by this definition and so is the BB84 quantum protocol. The question at the end of the previous paragraph was therefore: is there a quantum coin-flipping protocol whose bias is strictly less than 0.5?

The first such protocol was discovered in 2000 by Aharonov, Ta-Shma, Vazirani, and Yao [8], who proved that the bias of their protocol (ATVY) is at most  $\sqrt{2}-1 < 0.42$  (without any claim concerning the tightness of their bound). It was subsequently proven by Spekkens and Rudolph [9] that the ATVY protocol is better than its inventors had thought: its bias is in fact exactly  $\sqrt{2}/4 < 0.36$ . In the same paper, Spekkens and Rudolph gave an even simpler coin-flipping protocol that achieves the same bias as well as another one whose bias is merely  $(\sqrt{5}-1)/4 < 0.31$ . According to an earlier paper of theirs [10], that is the smallest bias possible for a coin-flipping protocol in which the quantum communication is limited to a single qubit.

Meanwhile, Ambainis [11] and, independently, Spekkens and Rudolph [10] discovered quantum coin-flipping protocols whose bias 0.25 is even smaller, but they require the transmission of a qutrit (or one qubit and two qutrits in the case of Spekkens and Rudolph). On the other hand, Kitaev [12] proved that no quantum coin-flipping protocol can have a bias below  $(\sqrt{2}-1)/2 \approx 0.21$ . Very recently, Chailloux and Kerenidis [13] announced a quantum coin-flipping protocol whose bias is arbitrarily close to Kitaev's bound, but it requires an unlimited number of rounds of interaction as it approaches this bound.

Despite the theoretical success of quantum coin-flipping protocols, compared to classical protocols, severe practical problems inherent to their implementation were discussed by Barrett and Massar [14], who argued that quantum coin flipping is problematic in any realistic scenario in which noise and loss can occur in the processing (preparation, transmission, and measurement) of quantum information. For this reason, they proposed random bit-string generation instead of *single-shot* coin flipping. However, this is not interesting from a quantum cryptographic perspective because the same goal can be achieved with purely classical means [15].

In a subsequent paper (NFPM) written in collaboration with Nguyen, Frison, and Phan Huy, Massar defined a *figure of merit* on which quantum coin-flipping protocols can outperform any possible classical protocol even in a realistic setting and they have verified their concept experimentally

[16]. Even though their protocol is not broken in the presence of loss, however, Alice can choose the outcome with near certainty in a realistic setting. We claim that, in order to be of practical use, a protocol should be *loss tolerant*, which we define as being completely impervious to loss of quantum information. In this sense, the NFPM protocol is not loss tolerant because its bias increases asymptotically towards 0.5 as losses become more and more severe, which is unavoidable in practice (with current technology) over increasing distance between Alice and Bob.

In this paper, we concentrate on this most likely source of imperfection in actual implementations, namely, *losses*. With the exception of the NFPM protocol mentioned above (which is not loss tolerant), all previously proposed quantum coin-flipping protocols become *completely* insecure even in the absence of noise as soon as the quantum channel between Alice and Bob is lossy. We introduce a loss-tolerant quantum coin-flipping protocol. We prove that our protocol is *fair* in the sense that either Alice or Bob can enjoy a bias of exactly 0.4 with an optimal cheating strategy, independently of the channel's transmission and other sources of losses, provided quantum information that is not lost is prepared, transmitted, and measured faithfully.

After this introduction, the structure of the paper is as follows. We begin in Sec. II with a review of the original 1984 quantum coin-flipping protocol of Bennett and Brassard [4] and we explain why it is completely vulnerable to a so-called EPR-attack. This is interesting not only for historical reasons but also because our loss-tolerant protocol follows the same template. Section III reviews perhaps the most famous of all quantum coin-flipping protocols, due to Ambainis [11], whose theoretical bias is 0.25. However, we demonstrate in Sec. IV that the security of that protocol is completely compromised in the presence of arbitrarily small channel loss. Moreover, we argue that this problem is inherent to the protocol in the sense that it cannot be repaired with small corrections. (The same would be true of the 0.25-bias qutrit-based coin-flipping protocol due to Spekkens and Rudolph [10] as well as of their optimal single-qubit protocol [9].) This is due to the notion of *conclusive measurements*, which we review in Sec. V. We show in Sec. VI how to combine the strengths of the original BB84 protocol with those of the ATVY protocol (which is not loss tolerant either in its published form) to finally achieve loss tolerance in quantum coin flipping and we analyze the security of our protocol. Conclusions and open problems are presented in Sec. VII.

## II. BB84 PROTOCOL

We review the original BB84 quantum coin-flipping protocol as well as the way it can be broken [4]. Here are the so-called BB84 states:

$$\left. \begin{aligned} |\psi_{0,0}\rangle &= |0\rangle \\ |\psi_{0,1}\rangle &= |1\rangle \end{aligned} \right\} a = 0,$$

$$\left. \begin{aligned} |\psi_{1,0}\rangle &= |+\rangle \\ |\psi_{1,1}\rangle &= |-\rangle \end{aligned} \right\} a = 1,$$

where  $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ . We say of  $|\psi_{a,x}\rangle$  that  $a$  is the *basis* and  $x$  is the *bit*. We define measurement bases

$$\mathcal{B}_a = \{|\psi_{a,0}\rangle, |\psi_{a,1}\rangle\} \quad (1)$$

for  $a \in \{0, 1\}$ . In the full BB84 quantum coin-flipping protocol [4], Alice would prepare and send Bob a large number of qubits, all in the same randomly chosen basis  $a$ . To emphasize the essential features of the protocol, however, we outline below a simplified version in which a single qubit is used.

(1) Alice prepares one of the four BB84 states  $|\psi_{a,x}\rangle$  with basis  $a$  and bit  $x$  chosen independently at random; she transmits that qubit to Bob.

(2) Bob chooses a random  $\hat{a} \in \{0, 1\}$  and measures the received qubit in basis  $\mathcal{B}_{\hat{a}}$ ; let Bob's result be  $\hat{x}$ .

(3) Bob sends a randomly chosen bit  $b$  to Alice.

(4) Alice reveals her original  $a$  and  $x$  to Bob.

(5) If  $a = \hat{a}$  and  $x \neq \hat{x}$ , Bob aborts the protocol, *calling Alice a cheater*; if  $a \neq \hat{a}$ , Bob has no way to verify Alice's honesty.

(6) If Bob did not abort the protocol, the outcome of the coin flip is  $a \oplus b$ , where " $\oplus$ " denotes the sum modulo 2 (also known as the "exclusive OR").

In this protocol, Bob cannot cheat at all. The only strategy for a cheating Bob would be to make an educated guess on Alice's choice of  $a$  before deciding on the  $b$  to send her at step 3 so as to bias the coin-flip outcome  $a \oplus b$ . However,  $a$  corresponds to Alice's random choice of basis. Hence, the state  $\rho_a$  received by Bob at step 1 is either  $\rho_0 = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$  or  $\rho_1 = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$ . It follows from the fact that  $\rho_0 = \rho_1$  that it is impossible for Bob to guess the value of  $a$  better than at random.

On the other hand, it is obvious that Alice can bias the protocol if she does not mind the risk of being called a cheater. The simplest approach is to be honest in the first step. When she receives  $b$  from Bob, the probability is 50% that she is happy with the outcome  $a \oplus b$ , in which case she proceeds honestly with the protocol. On the other hand, if she is unhappy with  $a \oplus b$ , she can lie on  $a$  at step 4 and send a random  $x$ . In that case, her probability of being caught is 25% since Bob chose  $\hat{a} \neq a$  with probability 50%, in which case he obtained  $\hat{x} \neq x$  also with probability 50%. All counted, this allows her to enjoy a 0.375 bias. A slightly more interesting cheat is for her to send state  $(\cos \frac{k\pi}{8})|0\rangle + (\sin \frac{k\pi}{8})|1\rangle$  for a random  $k \in \{1, 3, 5, 7\}$  at step 1 and declare the  $a$  that suits her wish (with the appropriately chosen  $x$ ) at step 4. This allows her to enjoy a bias of  $\frac{1}{2}\cos^2 \frac{\pi}{8} = (2 + \sqrt{2})/8 > 0.42$ , with a probability  $\frac{1}{2}\sin^2 \frac{\pi}{8} = (2 - \sqrt{2})/8 < 8\%$  of being called a cheater. It was to make the probability of undetected cheating exponentially small that the full BB84 protocol required the transmission and measurement of a large number of qubits [4].

A much more remarkable kind of cheating is possible for Alice, as explained in the same paper that introduced the BB84 protocol itself [4], which allows her to break the protocol completely (i.e., enjoy a 0.5 bias) with no fear of ever being caught. Let us say she wishes the outcome of the coin flip to be bit  $c$ . Instead of sending a legitimate BB84 state or

any other pure state at step 1, Alice sends half an EPR pair  $|\Psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$  to Bob and keeps the other half for herself. She waits until step 3, when she learns Bob's choice of  $b$ , to measure in basis  $a = c \oplus b$  the half she had kept; let  $x$  be her measurement outcome. This tells her that Bob has obtained (or will obtain if he has not yet measured)  $\hat{x} = 1 \oplus x$  in case he has measured (or will measure) in basis  $a$ . (Alice does not care about the value of  $\hat{x}$  if Bob chooses to measure in the other basis.) Hence, she can always obtain her desired outcome by sending those  $a$  and  $\hat{x}$  to Bob in step 4.

As subsequently discovered independently by Mayers [17] and by Lo and Chau [18] in the context of quantum bit commitment, this kind of cheating is *always* possible for Alice in any quantum coin-flipping protocol that has the structure of the BB84 protocol, regardless of the actual set of quantum states, whenever the density matrices used to signal  $a=0$  or  $a=1$  at step 1 are identical ( $\rho_0 = \rho_1$ ). This is due to the striking quantum process known as "remote steering," discovered by Schrödinger [19] as early as 1936 and better known as the Hughston-Jozsa-Wootters (HJW) theorem [20]. (See Ref. [21] for an entertaining history of this theorem.) Note that the remote steering attack works just as well if Bob postpones his measurement until after Alice reveals  $a$  and  $x$  and *almost* just as well if  $\rho_0$  and  $\rho_1$ , although different, are exponentially indistinguishable [22]. This last remark caused the demise of the bit commitment scheme proposed in Ref. [23], and we shall henceforth not differentiate between density matrices that are equal and those that are merely exponentially indistinguishable.

### III. AMBAINIS' PROTOCOL

In order to escape the remote steering attack, Aharonov *et al.* [8] introduced a coin-flipping protocol (ATVY) in which  $\rho_0 \neq \rho_1$ . To reduce even further Alice's possible bias, they shuffled the order of steps 2, 3, and 4 so that Bob delays his measurement of Alice's supplied state until after she tells him what she claims to have sent. In this way, he can measure systematically in the declared basis rather than having to measure in a randomly chosen basis  $\hat{a}$  whose outcome had probability 50% of being useless. This resulted in a quantum coin-flipping protocol with bias  $\sqrt{2}/4 < 0.36$ , as subsequently proven by Spekkens and Rudolph [10].

To achieve his smaller bias of 0.25, Ambainis used the following states on qutrits (rather than on qubits):

$$\left. \begin{aligned} |\phi_{0,0}\rangle &= \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \\ |\phi_{0,1}\rangle &= \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle \end{aligned} \right\} a=0, \quad \left. \begin{aligned} |\phi_{1,0}\rangle &= \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|2\rangle \\ |\phi_{1,1}\rangle &= \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|2\rangle \end{aligned} \right\} a=1.$$

Again we say of  $|\phi_{a,x}\rangle$  that  $a$  is the basis and  $x$  is the bit. This time, we define measurement bases



$$\mathcal{B}'_a = \{|\phi_{a,0}\rangle, |\phi_{a,1}\rangle, |2-a\rangle\}$$

for  $a \in \{0, 1\}$ . Here is Ambainis' protocol.

(1) Alice prepares one of the four Ambainis states  $|\phi_{a,x}\rangle$  with basis  $a$  and bit  $x$  chosen independently at random; she transmits that qutrit to Bob, who stores it in his quantum memory.

(2) Bob sends a randomly chosen bit  $b$  to Alice.

(3) Alice reveals her original  $a$  and  $x$  to Bob.

(4) Bob takes Alice's qutrit out of quantum memory and measures it in basis  $\mathcal{B}'_a$ ; let Bob's result be  $\hat{x}$ .

(5) If  $x \neq \hat{x}$ , Bob aborts the protocol, calling Alice a cheater.

(6) If Bob did not abort the protocol, the outcome of the coin flip is  $a \oplus b$ .

Alice cannot use remote steering to gain complete control over the outcome of the coin flip because the density matrices that could be received by Bob at step 1, corresponding to her choice  $a$  of basis,

$$\begin{aligned} \rho_0 &= \frac{1}{2}|\phi_{0,0}\rangle\langle\phi_{0,0}| + \frac{1}{2}|\phi_{0,1}\rangle\langle\phi_{0,1}| = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \rho_1 &= \frac{1}{2}|\phi_{1,0}\rangle\langle\phi_{1,0}| + \frac{1}{2}|\phi_{1,1}\rangle\langle\phi_{1,1}| = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1/2 \end{pmatrix}, \end{aligned} \quad (2)$$

are distinct. It is easy to see that Alice can enjoy a 0.25 bias if she sends state  $(2|0\rangle \pm |1\rangle \pm |2\rangle)/\sqrt{6}$  at step 1 and declares  $a$  and  $x$  appropriately at step 3. The proof that this is Alice's optimal cheating strategy is nontrivial but has been worked out in detail by Ambainis [11]. On the other hand, the fact that  $\rho_0 \neq \rho_1$  makes it possible for Bob to cheat by measuring Alice's qutrit before step 2 in order to learn information about  $a$  and bias his choice of  $b$  accordingly. The most obvious strategy for Bob is to measure Alice's qutrit in the computational basis  $\{|0\rangle, |1\rangle, |2\rangle\}$ , which allows him to enjoy a 0.25 bias as well. The fact that this is Bob's optimal cheating strategy follows directly from Helstrom's optimal measurement theory [24], which we review in Sec. V.

For completeness, we mention that the independently discovered 0.25-bias quantum coin-flipping protocol of Spekkens and Rudolph [10] requires Alice to choose a random bit  $a$ , prepare a one-qubit-and-two-qutrit entangled state

$$|\Gamma_a\rangle = |a\rangle \otimes \left( \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|a+1, a+1\rangle \right),$$

and send Bob one of the two qutrits in step 1. Step 2 is the same as in Ambainis' protocol. In step 3, Alice sends the qubit and the other qutrit to Bob, which allows him to verify her honesty in step 4 by performing a positive-operator-valued measure (POVM) on the combined Alice-provided state whose outcome is either  $|\Gamma_0\rangle$ ,  $|\Gamma_1\rangle$  or "Alice cheated." Again, the result of the coin flip is  $a \oplus b$  provided Bob did not catch Alice cheating. The essential (but not sufficient) reason why this protocol gives the same bias as Ambainis' is that the density matrix  $\rho_a$  received by Bob at step 1 is exactly the same in both protocols, as given by Eq. (2).

#### IV. PRACTICAL VULNERABILITY

Even though Ambainis' analysis of his protocol is mathematically impeccable, there is a practical problem that cannot be neglected if one is ever to implement such protocols in real life: there will be unavoidable losses in the quantum channel between Alice and Bob. This is true in particular if photons are used to carry quantum information and if the quantum channel is an optical fiber. Further losses are to be expected in Bob's quantum memory and detection apparatus. The situation is even worse for the 0.25-bias protocol of Spekkens and Rudolph because it requires both Alice and Bob to have quantum memory since they must store one qutrit each between steps 1 and 3 (there is no real need for Alice to store also the qubit since it contains only classical information). Please remember that one of the main appeals of quantum cryptography, from its very beginnings [4], has been to offer protocols *that can be implemented with current technology* yet remain secure against any potentially future attack so long as quantum mechanics is not violated.

It follows that there is a possibility<sup>1</sup> when Bob tries to measure Alice's qutrit at step 4 of Ambainis' protocol (or the one-qubit-two-qutrit system in the protocol of Spekkens and Rudolph) that he does not register anything *even though both Alice and Bob have been entirely honest*. How should Bob react in this case? He could hardly call Alice a cheater if the most likely cause for loss is in his own detectors. There seem to be only two reasonable responses from Bob, as pointed out already by Barrett and Massar [14] concerning a quantum coin-flipping protocol inspired by the quantum gambling protocol of Hwang *et al.* [25]: Bob can (1) accept Alice's declared  $a$  and  $x$  on faith or (2) request Alice to restart the coin-flipping protocol from scratch. But both these "solutions" are unacceptable.

If Alice knows that Bob will believe her on faith in case he gets no detection signal, she can totally bias the coin flip with the most maddeningly simple cheating strategy: she does *nothing* at all during step 1 and lets Bob "store" the empty signal.<sup>2</sup> After having received Bob's choice of  $b$ , she is then free to "reveal" whichever  $a$  would produce her desired outcome  $a \oplus b$ . Since Bob's measurement of the empty signal will yield nothing at step 4, he has no other choice but to believe Alice on faith.

On the other hand, if Alice and Bob have agreed to restart the coin-flipping protocol in case Bob does not detect anything at step 4, it is Bob who can totally bias the coin flip without any need to manipulate quantum states. Whenever he receives Alice's qutrit, he does nothing at all with it and sends his random choice of  $b$ . If Alice reveals  $a$  at step 3 so that he is happy with outcome  $a \oplus b$ , with probability 50%, Bob pretends to measure Alice's long-lost qutrit and claims to be satisfied with her honesty. But if he is not happy with

<sup>1</sup>Actually, with current technology, it is more than a mere "possibility:" this will occur in the overwhelming majority of cases.

<sup>2</sup>It would be technologically very challenging for Bob to perform a so-called "non-demolition measurement," which would in principle allow him to confirm the presence of the qutrit without disturbing its state.

the outcome, Bob simply tells Alice that he has not registered anything and requests a new instance of the protocol. This continues until the outcome is to Bob's liking. This cheating strategy cannot be detected by Alice whenever the expected probability  $p$  of registering an outcome when both Alice and Bob are honest is at most 50%, provided Bob asks to restart the protocol with probability  $1-2p$  even when progressing to step 6 would have produced his desired outcome.

Unless Bob has the technological ability to make sure he received a quantum state from Alice at step 1 without disturbing it and because Alice will not accept to restart the protocol after she has revealed  $a$  and  $x$ , he has only one line of defence against her "send nothing" cheating strategy: he must measure Alice's signal immediately upon reception and be allowed to ask her to restart the protocol from scratch (with an independent random choice of state) until he actually registers a measurement outcome. This means that we must revert to the original BB84 template in which Bob measures before Alice reveals  $a$  and  $x$ .

As explained at the beginning of Sec. III, this will make it easier for a cheating Alice to escape detection since Bob's measurement can no longer depend on her claimed state. However, provided Bob chooses his measurement basis at random, he will carry out with probability 50% the same measurement he would have performed in Ambainis' original protocol. As we prove in Sec. VI A, it follows that Bob's probability of catching Alice's eventual cheating is reduced but not by more than a factor of 2, compared to the original protocol. Hence, Alice's bias is at most 0.375 rather than 0.25. The important point is that this bias remains below 0.5.

Nevertheless, this modification in Ambainis' protocol, which is made necessary by practical considerations (with current technology), reopens the door for Bob to completely break the revised protocol. When Alice tells him she transmitted a qutrit, Bob measures it immediately in the computational basis  $\{|0\rangle, |1\rangle, |2\rangle\}$ . If he obtains either  $|1\rangle$  or  $|2\rangle$ , Bob *knows* Alice's choice of  $a$ . This allows him to choose  $b$  so that  $a \oplus b$  suits his desired outcome. On the other hand, if Bob either obtains  $|0\rangle$  or if he does not register anything, then he tells Alice that the transmission has been unsuccessful and he requests another qutrit from her. In effect, the protocol will proceed to the step in which Bob sends his choice of  $b$  to Alice *only* when he already knows Alice's earlier choice of  $a$ . This means that Bob enjoys a bias of 0.5 and the protocol is completely broken. To camouflage his chicanery (in case Alice might wonder why the measurement does not succeed more often), Bob can pretend at the outset that his detectors are half as efficient as they really are. Alternatively, he could surreptitiously replace the quantum channel that links him to Alice with a sufficiently better one.

This fatal flaw in any practical implementation of Ambainis' protocol comes from one simple consideration. Even though the mixed states  $\rho_0$  and  $\rho_1$  [see Eq. (2)] used at step 1 by Alice to partially commit to either  $a=0$  or  $a=1$ , respectively, are nonorthogonal, hence they cannot be distinguished with certainty by Bob all the time, *they can be distinguished conclusively with positive probability*. After reviewing below the notion of conclusive measurements, we introduce our protocol in Sec. VI and prove that its bias is exactly 0.4 even when arbitrarily severe losses are taken into account.

## V. DIFFERENT TYPES OF MEASUREMENTS

Consider two nonorthogonal density matrices  $\rho_0$  and  $\rho_1$  (they could be pure states). There are several figures of merit in measurements that attempt to distinguish them [22]. Helstrom studied the optimal measurement to output a guess that minimizes the error probability [24]. Assuming both  $\rho_0$  and  $\rho_1$  were equally likely *a priori*, Helstrom's measurement outputs the correct guess with probability

$$\frac{1}{2} + \frac{1}{2} D(\rho_0, \rho_1), \quad (3)$$

where

$$D(\rho_0, \rho_1) = \frac{1}{2} \text{Tr} |\rho_0 - \rho_1| \quad (4)$$

is the *trace distance* between  $\rho_0$  and  $\rho_1$ , "Tr" denotes the trace, and  $|A| = \sqrt{A^\dagger A}$ . In particular, if  $A$  is a diagonal real matrix, then  $|A|_{ij} = |A_{ij}|$ .

When the spans of  $\rho_0$  and  $\rho_1$  are distinct, there exists another type of measurement, known as conclusive measurement or *unambiguous state discrimination* (USD) [26–28]. These measurements have three possible outcomes, "0," "1," and "?," the latter of which being called the *inconclusive outcome*. Whenever outcome  $a \in \{0, 1\}$  is obtained, it is guaranteed that the measured state was indeed  $\rho_a$  (assuming of course that it was either  $\rho_0$  or  $\rho_1$  and that there were no experimental errors). Furthermore, the probability of obtaining a conclusive outcome (not "?") must be strictly positive for either input state.

A well-known example of conclusive measurement can distinguish between  $|0\rangle$  and  $|+\rangle$  with conclusive outcome probability  $1 - 1/\sqrt{2} > 29\%$ . More to the point of our paper, the obvious measurement in computational basis  $\{|0\rangle, |1\rangle, |2\rangle\}$  distinguishes between Ambainis'  $\rho_0$  and  $\rho_1$  [Eq. (2)] with conclusive probability 50%. In general, any coin-flipping protocol that follows the template of BB84 is vulnerable to the attack described in Sec. IV when a conclusive measurement exists between the corresponding  $\rho_0$  and  $\rho_1$ .

It is tempting to think that this line of attack will not apply when conclusive measurements do not exist. However, this is not necessarily the case. Barnett and co-workers [29,30] studied "maximum confidence quantum measurements" (MCQM), which somehow interpolate between Helstrom and conclusive measurements. Like conclusive measurements, MCQMs have some probability  $p < 1$  of yielding the "?" inconclusive outcome. However, when the outcome is either "0" or "1," it is correct with probability at least  $q > 0$ . Helstrom's measurement maximizes  $q$  subject to  $p=0$ , whereas conclusive measurements (when they exist) minimize  $p$  subject to  $q=1$ . There can be a trade-off between those two probabilities in the sense that it is sometimes possible to achieve an increase in  $q$  (compared to the accuracy of Helstrom's measurement) by tolerating a larger  $p$ .

The relevance of MCQMs is that Bob could use them to increase his coin-flip bias at the cost of increasing his probability of asking Alice to restart the protocol (pretending he has not registered an outcome). If MCQMs exist for  $q$  arbitrary,

trarily close to 1, Bob can come as near as he wants to choosing the coin-flip outcome, provided Alice has enough patience (naïvety?) to accept restarting the protocol indefinitely until Bob tells her that he is ready to continue.

To demonstrate that this is a legitimate worry, consider an (admittedly contrived) quantum coin-flipping protocol based on Ambainis' states (see the beginning of Sec. III), except that we add  $|\phi_{0,2}\rangle=|2\rangle$  and  $|\phi_{1,2}\rangle=|1\rangle$ . In step 1, Alice chooses basis  $a \in \{0,1\}$  at random with uniform probability, but  $x \in \{0,1,2\}$  is chosen so that  $\text{Prob}(x=0) = \text{Prob}(x=1) = 49\%$ , whereas  $\text{Prob}(x=2) = 2\%$ . The density matrix received by Bob would be either

$$\rho'_0 = \begin{pmatrix} 0.49 & 0 & 0 \\ 0 & 0.49 & 0 \\ 0 & 0 & 0.02 \end{pmatrix} \text{ or } \rho'_1 = \begin{pmatrix} 0.49 & 0 & 0 \\ 0 & 0.02 & 0 \\ 0 & 0 & 0.49 \end{pmatrix}.$$

These two mixed states cannot be distinguished conclusively. Nevertheless, a measurement in the computational basis yields either  $|0\rangle$ , which is interpreted as the inconclusive outcome “?”, or it yields either  $|1\rangle$  or  $|2\rangle$ , which are interpreted as either  $\rho'_0$  or  $\rho'_1$ , respectively. This MCQC is inconclusive with probability  $p=49\%$ . When it is not inconclusive, however, the verdict is correct with probability  $q=0.49/0.51 > 96\%$ . This is much better than Helstrom's measurement, which would always give an answer but be correct only with probability 73.5% since  $D(\rho'_0, \rho'_1) = 0.47$ . Therefore, a quantum coin-flipping protocol that uses these states would allow Bob to enjoy a maximal theoretical bias of 0.235 if we did not take losses into account. However, Bob's bias becomes larger than 0.46 if Alice agrees to restart the protocol whenever he claims to have not registered an outcome, which he would do whenever his measurement outcome is the inconclusive  $|0\rangle$ .

## VI. LOSS-TOLERANT PROTOCOL

Now, we introduce our quantum coin-flipping protocol and we prove that its bias is exactly 0.4, regardless of the extent of losses that are unavoidable with current technology. For this, we use the states introduced in the ATVY protocol [8] but revert to the original BB84 template [4]. Consider the states

$$\left. \begin{aligned} |\varphi_{0,0}\rangle &= \alpha|0\rangle + \beta|1\rangle \\ |\varphi_{1,0}\rangle &= \alpha|0\rangle - \beta|1\rangle \end{aligned} \right\} x=0,$$

$$\left. \begin{aligned} |\varphi_{0,1}\rangle &= \beta|0\rangle - \alpha|1\rangle \\ |\varphi_{1,1}\rangle &= \beta|0\rangle + \alpha|1\rangle \end{aligned} \right\} x=1,$$

where  $\alpha$  and  $\beta$  are real numbers such that  $1 > \alpha > \beta > 0$  and  $\alpha^2 + \beta^2 = 1$ . See Fig. 1. As always, we say of  $|\varphi_{a,x}\rangle$  that  $a$  is the basis and  $x$  is the bit. We define measurement bases *mutatis mutandis* as we had done in Eq. (1) for the BB84 states:

$$\mathcal{B}_a'' = \{|\varphi_{a,0}\rangle, |\varphi_{a,1}\rangle\} \quad (5)$$

for  $a \in \{0,1\}$ . We shall soon see why we have regrouped the states according to the value of  $x$  rather than that of  $a$  as we had done previously. Here is our loss-tolerant quantum coin-flipping protocol.

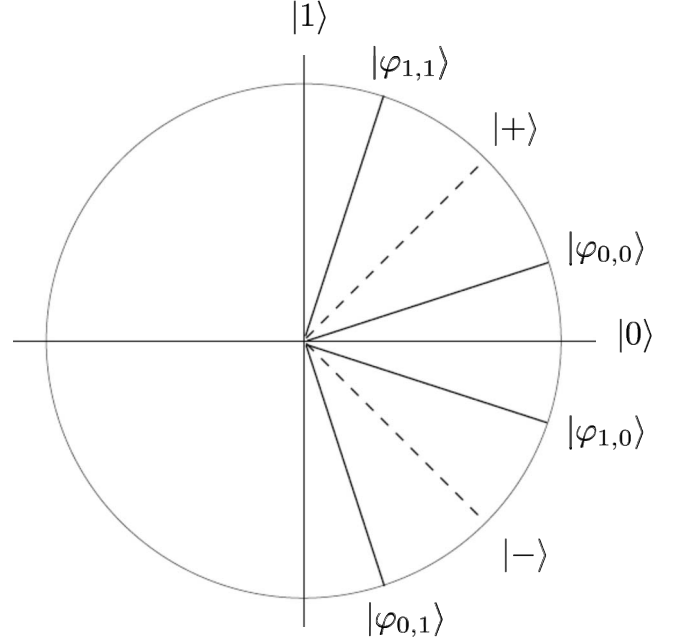


FIG. 1. Quantum states used in loss-tolerant protocol. States such as  $|\varphi_{0,0}\rangle = \alpha|0\rangle + \beta|1\rangle$  are represented in the real Cartesian plane by a line between the origin and point  $(\alpha, \beta)$ .

(1) Alice prepares one of the four states  $|\varphi_{a,x}\rangle$  with basis  $a$  and bit  $x$  chosen independently at random; she transmits that qubit to Bob.

(2) Bob chooses a random  $\hat{a} \in \{0,1\}$  and measures the received qubit in basis  $\mathcal{B}_{\hat{a}}''$ . If his apparatus does not register an outcome, he requests Alice to start over at step 1; otherwise, let the measurement result be  $\hat{x}$ .

(3) Bob sends a randomly chosen bit  $b$  to Alice.

(4) Alice reveals her original  $a$  and  $x$  to Bob.

(5) If  $a = \hat{a}$  and  $x \neq \hat{x}$ , Bob aborts the protocol, calling Alice a cheater; if  $a \neq \hat{a}$ , Bob has no way to verify Alice's honesty.

(6) If Bob did not abort the protocol, the outcome of the coin flip is  $x \oplus b$ .

There are three differences, compared to the original BB84 protocol described in Sec. II: (1) in addition to having been globally rotated to facilitate subsequent analysis, the states used are more general in the sense that bases  $\mathcal{B}_0''$  and  $\mathcal{B}_1''$  need not be mutually unbiased; (2) step 2 allows Bob to ask Alice to restart the protocol in case his measurement apparatus fails to register an outcome; and (3) the final coin-flip result is  $x \oplus b$  rather than  $a \oplus b$ .

The first modification will allow us to fine-tune the protocol in Sec. VI C. The second modification makes it possible for the protocol to be loss tolerant as we have explained in Sec. IV. The effect of the third modification, which corresponds to the main original contribution of the ATVY protocol, is that two *distinct* density matrices  $\rho_0$  and  $\rho_1$  (see below) are now used by Alice to partially commit to either  $x=0$  or  $x=1$  by the transmission of  $|\varphi_{a,x}\rangle$  at step 1 with a randomly chosen  $a$ :



$$\begin{aligned}\varrho_0 &= \frac{1}{2}|\varphi_{0,0}\rangle\langle\varphi_{0,0}| + \frac{1}{2}|\varphi_{1,0}\rangle\langle\varphi_{1,0}| = \begin{pmatrix} \alpha^2 & 0 \\ 0 & \beta^2 \end{pmatrix}, \\ \varrho_1 &= \frac{1}{2}|\varphi_{0,1}\rangle\langle\varphi_{0,1}| + \frac{1}{2}|\varphi_{1,1}\rangle\langle\varphi_{1,1}| = \begin{pmatrix} \beta^2 & 0 \\ 0 & \alpha^2 \end{pmatrix}.\end{aligned}\quad (6)$$

In sharp contrast, the intuition behind the original BB84 protocol was for Alice to “commit” to either  $a=0$  or  $a=1$  with a randomly chosen  $x$  but that was doomed by Schrödinger’s remote steering process because the corresponding density matrices were equal.

Furthermore, we shall see in Sec. VI B that, contrary to the mixed states  $\rho_0$  and  $\rho_1$  used in Ambainis’ protocol [see Eq. (2)], this time  $\varrho_0$  and  $\varrho_1$  cannot be distinguished conclusively nor even by a maximum confidence quantum measurement better than Helstrom’s measurement, which is the key to loss tolerance.

Although we have presented our protocol as a modification of the original BB84 protocol, it is useful for analysis purposes to contrast it also with the ATVY protocol. The key difference between the ATVY protocol and ours is that they chose to introduce a new template (used in most subsequent protocols such as Ambainis’) in which Bob stores the quantum state sent by Alice at step 1 in order to delay measurement until Alice has given him its classical description. The intention was to make it harder for Alice to cheat because Bob would know to measure the state in basis  $\mathcal{B}_a''$  rather than having to choose a random basis  $\mathcal{B}_a'$ . It was this feature that made their protocol incapable of tolerating channel loss unless Bob has the technological ability to detect if a signal has been received by Alice without otherwise disturbing its state.

Next, we prove that the maximum biases that Alice and Bob can enjoy are

$$\varepsilon_A = (1 + 2\alpha\beta)/4 \quad \text{and} \quad \varepsilon_B = \alpha^2 - 1/2,$$

respectively, and we give explicit cheating strategies to achieve those biases. We conclude that the choice of  $\alpha$  and  $\beta$  that makes those two biases equal corresponds to a fair protocol whose bias is 0.4.

#### A. Alice’s optimal cheating strategy

It would be relatively easy to determine Alice’s optimal cheating strategy were she restricted to sending some pure state to Bob in the first step of the protocol. However, we have learned from the demise of the original BB84 protocol [4] that it might be to her dishonest advantage to prepare an entangled state, send one qubit from it to Bob at step 1, and wait until Bob’s announcement of  $b$  at step 3 to measure in the most informative way what she had kept. This could increase her chances of deciding on her best choice of  $a$  and  $x$  to “reveal” at step 4 in order to maximize her probability of successfully biasing the coin flip. It is significantly more complicated to take all possible such strategies into account. Fortunately, most of the work has already been done by Spekkens and Rudolph [9] in their thorough analysis of the ATVY protocol.

Let us begin our analysis by briefly pretending that Bob does not measure the state Alice has sent him until after she

reveals her choice of basis and bit and that he measures it in that basis. As we have already pointed out, this becomes the ATVY protocol. It follows directly from the analysis of Spekkens and Rudolph [choosing what they call  $\theta$  in their Eq. (20) so that  $\cos \theta = \alpha$  and  $\sin \theta = \beta$ ] that any cheating strategy Alice may deploy gives her bias

$$\varepsilon'_A \leq \frac{\sin 2\theta}{2} = \sin \theta \cos \theta = \alpha\beta. \quad (7)$$

Now, to analyze our protocol, we must take into account the fact that we require Bob to measure Alice’s state *before* she tells him what she claims to have sent. (This is how we make our protocol loss tolerant.) This difference makes it easier for Alice to cheat because Bob’s measurement cannot be chosen to maximize his probability of discriminating between her claimed state and any other state that she might have sent instead. However, as we show below, this modification does not reduce Bob’s probability of catching Alice cheating by more than a factor of 2.

Recall that Bob is required by step 3 of the protocol to send a *random* choice of  $b$ , which must be uncorrelated with his former choice  $\hat{a}$  of measurement and its outcome  $\hat{x}$ . The crucial observation is that the randomness of  $b$  deprives Alice from any information she might otherwise have obtained from Bob concerning  $\hat{a}$  and  $\hat{x}$ . (The importance of this observation is illustrated in Sec. VI D.) It follows that her choice of which  $a$  and  $x$  to declare at step 4 cannot depend on what has happened at Bob’s after she transmitted her quantum state at step 1. In particular,  $\hat{a}=a$  with probability 50% since  $\hat{a}$  is chosen at random by an honest Bob, in which case he has chosen by chance at step 2 precisely the measurement he would have performed in the ATVY protocol, had he been allowed to wait until Alice’s announcement of  $a$  and  $x$  before deciding on his measurement.

The above implies that any cheating strategy that Alice might deploy against our protocol translates into an *identical* cheating strategy against the ATVY protocol, possibly with a different success probability, since it makes no measurable difference to Alice whether Bob measures before (as in our protocol) or after (as in the ATVY protocol) she has to declare  $a$  and  $x$ . Now, consider an arbitrary cheating strategy on the part of Alice and let  $\varepsilon_A$  (respectively,  $\varepsilon'_A$ ) be the bias that she enjoys with this strategy against our protocol (respectively, the ATVY protocol).

Consider an arbitrary run of our protocol, when Alice uses this cheating strategy. With probability 50%, independently from anything else, Bob randomly chooses the same measurement he would have performed in the ATVY protocol, in which case Alice succeeds with probability  $\frac{1}{2} + \varepsilon'_A$ . In the other 50% of the cases, Bob cannot verify the state Alice claims to have sent and the probability that Alice succeeds is at most 1. All counted, the success probability of Alice in our protocol is

$$\frac{1}{2} + \varepsilon_A \leq \frac{1}{2} \left( \frac{1}{2} + \varepsilon'_A \right) + \frac{1}{2} \times 1 = \frac{3}{4} + \frac{1}{2} \varepsilon'_A.$$

It follows that

$$\varepsilon_A \leq \frac{1 + 2\varepsilon'_A}{4} \leq \frac{1 + 2\alpha\beta}{4}, \quad (8)$$

where the last inequality follows from Eq. (7).

We now proceed to show that this bound can be saturated. Surprisingly (when we recall the demise of the BB84 quantum coin-flipping protocol), the optimal cheating strategy for Alice does not require her to prepare an entangled state from which she would send one qubit at step 1. Instead, it suffices for her to send either

$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

or

$$|-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

to Bob. Let us say she sent  $|+\rangle$  (the other case is similar) and received bit  $b$  from Bob at step 3. If her desired outcome is  $c$ , she sets  $x = c \oplus b$  and claims at step 4 to have sent  $|\varphi_{x,x}\rangle$  at step 1. With probability 50%, Bob had already chosen  $\hat{a} \neq x$ , in which case he cannot catch Alice cheating. With complementary probability 50%, he had chosen  $\hat{a} = x$ , in which case Alice escapes detection with probability

$$|\langle + | \varphi_{x,x} \rangle|^2 = \left( \frac{1}{\sqrt{2}}\alpha + \frac{1}{\sqrt{2}}\beta \right)^2 = \frac{(\alpha + \beta)^2}{2} = \frac{1}{2} + \alpha\beta$$

(the last equality is because  $\alpha^2 + \beta^2 = 1$ ). Putting it all together, Alice obtains her desired outcome with probability

$$\frac{1}{2} + \frac{1}{2} \left( \frac{1}{2} + \alpha\beta \right) = \frac{3 + 2\alpha\beta}{4}$$

and her bias is therefore

$$\varepsilon_A = \frac{3 + 2\alpha\beta}{4} - \frac{1}{2} = \frac{1 + 2\alpha\beta}{4},$$

which saturates the bound given in Eq. (8).

### B. Bob's optimal cheating strategy

In the analysis of Bob's bias in quantum coin-flipping protocols, it is usual to apply Eqs. (3) and (4) to compute the trace distance  $D(\varrho_0, \varrho_1) = \alpha^2 - \beta^2 = 2\alpha^2 - 1$  in order to determine that the optimal Helstrom measurement that Bob could perform to best guess Alice's committed bit (here,  $x$ ) gives him the correct answer with probability  $\alpha^2$ . From this, we would "normally" conclude that Bob's maximal bias is  $\alpha^2 - 1/2$ .

As we have seen, however, this approach is not appropriate in our context because it does not take into account the eventual possibility for Bob to increase his bias by making conclusive or maximum confidence quantum measurements on the state sent by Alice at step 1 so that he could ask her to restart the protocol whenever he is not sufficiently satisfied with the probability that his guess of  $x$  be correct.

Fortunately, the analysis of Bob's optimal cheating strategy is straightforward in this case. For either value of

$x \in \{0, 1\}$  that Alice may have chosen in step 1, Eq. (6) shows the mixed state  $\varrho_x$  that Bob would receive from her. The key observation is that, algebraically speaking,  $\varrho_0 = \alpha^2|0\rangle\langle 0| + \beta^2|1\rangle\langle 1|$  and  $\varrho_1 = \beta^2|0\rangle\langle 0| + \alpha^2|1\rangle\langle 1|$ . It follows that, from a cheating Bob's perspective, whose only concern is in guessing  $x$  as best as possible (including the possibility of asking Alice to restart the protocol if he is not satisfied with his own confidence in his guess), this is *strictly identical* to an alternative scenario in which Alice would have sent  $|0\rangle$  with probability  $\alpha^2$  or  $|1\rangle$  with probability  $\beta^2$  in case  $x=0$ , and vice versa in case  $x=1$ . In other words, this is exactly as if Alice had communicated her choice of  $x$  to Bob as the *purely classical* signal  $|x\rangle$  through a binary symmetric channel with error probability  $\beta^2$ .

Seen this way, there is nothing quantum about the situation. Hence, a complete measurement in the computational basis of the state received by Alice provides Bob with an optimal cheating strategy since it yields *all* the information classically available in the "quantum" signal. Not surprisingly, this is indeed precisely Helstrom's measurement. In particular, the outcome of this measurement does not give Bob any indication on whether or not it would be to his advantage to ask Alice to restart the protocol. (Note that it is not simply because Alice's signal can be thought of as classical that there cannot be a conclusive or a maximum confidence measurement; in particular, *erasure channels* provide the classical equivalent to quantum conclusive measurements—but binary symmetric channels do not.)

To summarize, Bob's optimal cheating strategy is to measure Alice's qubit in the computational basis in order to learn the value of  $x$  with error probability  $\beta^2$ . This allows him to choose  $b$  appropriately and obtain his desired outcome  $x \oplus b$  with complementary probability  $\alpha^2$ , thus enjoying his optimal bias  $\varepsilon_B = \alpha^2 - 1/2$ .

### C. Fair protocol

In this subsection, we find the value of  $\alpha$  so that the bias either player can achieve by cheating is the same. We simply need to fulfill the condition

$$\varepsilon_A = \varepsilon_B,$$

which for our protocol amounts to

$$\frac{1 + 2\alpha\beta}{4} = \alpha^2 - \frac{1}{2},$$

subject to  $\alpha^2 + \beta^2 = 1$ . Solving this system yields

$$\alpha = \sqrt{0.9} \quad \text{and} \quad \beta = \sqrt{0.1},$$

as illustrated in Fig. 1. These values correspond to

$$\varepsilon_A = \varepsilon_B = 0.4,$$

which defines a fair loss-tolerant quantum coin-flipping protocol whose bias is 0.4. In other words, either player can obtain a desired outcome with probability 90% by using an optimal cheating strategy, provided of course the other player is honest. But what if they both try to cheat?



### D. Cunning games

An interesting phenomenon is illustrated if we take the unusual step of considering the case when both Alice and Bob cheat. The most obvious example of double cheating occurs when Alice is convinced that Bob is so greedy that he will try his best to control the coin flip. In this case, Alice can send an honest quantum state in step 1. To maximize his chances of guessing Alice's choice of  $x$ , Bob measures the state in the computational basis, as we have seen. This allows him to send Alice a value of  $b$  that produces his desired outcome with probability 90% (assuming they use the fair version of the protocol). But at the same time, he has lost his ability to verify the honesty of Alice since he did not measure her qubit in a legitimate basis. Suspecting this, Alice is free to claim whatever suits her best at step 4, probably lying about  $x$ , and Bob cannot look her in the face and call her a cheater.

For an amusing variation on this theme, consider what happens if both parties attempt to perform their optimal cheating strategies simultaneously. In this case, Alice sends either  $|+\rangle$  or  $|-\rangle$ , which Bob measures in basis  $\{|0\rangle, |1\rangle\}$ . When Bob attempts influencing the coin flip by choosing  $b$  as a function of his measurement outcome, little does he suspect that his choice is in fact totally random: ironically, Bob follows the honest protocol at step 3.

For a more subtle example, consider a scenario according to which Bob is Alice's young son. They agree to flip a quantum coin to determine who will decide on the film to be seen tonight: Alice will have the choice if the outcome is 1 and Bob if it is 0. Alice knows that her little rascal will do his best to cheat and win but also that he will follow step 2 properly (rather than measuring in the computational basis) because he will not want to relinquish his possibility to catch his mother cheating at verification step 5—or so he thinks. Hence, Bob's cheating will consist in sending  $b=\hat{x}$  at step 3 instead of choosing  $b$  at random. After a calculation, we find that this gives him probability 82% that  $b=x$ , thus producing his winning outcome  $x \oplus b=0$ .

Unbeknownst to Bob, however, his mother wants him to win, but she does not wish him to know this for fear of undermining her authority. Assuming she knows her son as well as she thinks, she proceeds as follows. At step 1, she honestly sends some random state  $|\varphi_{a,x}\rangle$ . When she receives Bob's choice of  $b$  at step 3, there are two possibilities. If  $b=x$ , she can relax and continue honestly since Bob wins, according to her wish. On the other hand, if  $b \neq x$ , then Alice can deduce that Bob used the wrong basis in his measurement:  $\hat{a} \neq a$ . This allows her to "reveal"  $a$  together with any  $\tilde{x}$  of her choice at step 4 since she knows (or suspects) that Bob will be unable to call her a cheater. According to our story, she chooses  $\tilde{x}=b$  to make Bob happy.

Admittedly, the above tale is unlikely at best. Nevertheless, it serves the purpose of demonstrating the importance in our proof of security against Alice (Sec. VI A) that Bob's bit  $b$  be chosen randomly despite the fact that he has significant information about Alice's  $x$  after an honest measurement at step 2. Indeed, our proof relied in a crucial way on the fact that Alice would have no information on the measurement basis used by Bob. Our tale shows how wrong and damaging

it would have been had the protocol asked an *honest* Bob to send a  $b$  correlated to his measurement outcome.

### E. Side channels

We have seen that Bob could exploit the loss of quantum information to cheat in previous protocols, whereas our protocol is loss tolerant. Nevertheless, our protocol could become susceptible to loss if it were implemented with insufficient care. A *side channel* is any source of information that Bob could exploit about the quantum state sent by Alice above and beyond its theoretical definition as a pure qubit. For example, the protocol would be obviously insecure if the apparatus used by Alice to generate her quantum states produced each of the four legitimate states as photons of significantly different wavelengths or spatial position. These issues have been studied extensively in the context of quantum key distribution. See Ref. [31] for a compelling example of successful hacking of a commercially available apparatus.

An interesting example of side channel would occur in a careless implementation of our quantum coin-flipping protocol if Alice used an attenuated laser pulse to generate her states, as is done in most current implementations of quantum key distribution. The problem stems from the fact that one can distinguish conclusively between  $\varrho_0$  and  $\varrho_1$  [Eq. (6)] when a pulse consists of two (or more) identical ATVY states. For this, it suffices for Bob to measure one photon in basis  $\mathcal{B}_0''$  and the other in basis  $\mathcal{B}_1''$  [Eq. (5)]. If the two measurements produce the same outcome, this is necessarily the correct bit  $x$  encoded by Alice since one of the measurement has been performed in the correct basis. This occurs with probability  $(\alpha^2 - \beta^2)^2$ . With the fair states, Bob's probability of conclusive outcome is therefore a substantial 64% each time a pulse contains two photons, which means that this *implementation* of the protocol is completely insecure because Bob can request another state from Alice until this event occurs.

To make the case of attenuated laser pulses even worse, a perfectly natural measurement apparatus that an honest Bob could use in the implementation of our protocol [32] would produce such conclusive outcome with half the probability derived above, namely, 32% of the double-photon pulses. It follows that Bob, who started the protocol with pure intentions, might find it difficult to resist sliding to the dark side when his (honest) measurement apparatus reveals with certainty the value of  $x$  chosen by Alice. A practical solution to this problem is for Alice to use a source of entangled photons, as discussed in Ref. [32].

## VII. CONCLUSION

Quantum coin flipping is a cryptographic primitive that has been studied extensively. Several approaches have been considered since the very beginnings of quantum cryptography. However, previous protocols are either totally insecure [4] or they are highly sensitive to (if not completely broken by) technologically unavoidable losses of quantum information on the channel between the players or in their storage and detection apparatus [8–11,13,16]. In this paper, we intro-

duced a loss-tolerant quantum coin-flipping protocol, which means that it is completely impervious to such losses. We proved that our protocol is *fair* in the sense that both Alice and Bob have an optimal cheating strategy capable of producing their desired outcome with 90% probability of success (assuming the other player is honest) and we provided those strategies explicitly.

Even though our protocol can tolerate arbitrary loss of quantum information, it would fail in case of *noise* because it would be impossible for Bob to know, in case of a mismatch between his measurement outcome and Alice's claimed state, if that is due to a genuine error (on his part or Alice's) or to Alice's dishonesty. Recall that this may be unavoidable; indeed Barrett and Massar [14] argued that single-shot quantum coin-flipping protocols are problematic when both noise and loss can occur.

If secure single-shot quantum coin-flipping protocols are indeed impossible in the presence of noise, is there something useful that quantum coin flipping can do above and beyond anything classically possible? We already know that random bit-string generation [14] is not the answer given that it can be done classically [15]. We are currently investigating this issue [32] along a different line from that proposed in Ref. [16].

Another open question concerns the optimality of our protocol. Could there be a loss-tolerant quantum coin-flipping protocol whose bias is smaller than 0.4? Alternatively, can we prove that 0.4 is the smallest bias possible among all loss-tolerant quantum coin-flipping protocols? Finally, we mention that this paper was entirely concerned with the task known as strong coin flipping. There exists a similar task, *weak* coin flipping, in which only one outcome is favorable for Alice, the other outcome being the only one favorable for Bob. Protocols with arbitrarily small bias are known to exist

for weak coin flipping [33], in stark contrast with Kitaev's lower bound for the strong case [12]. Can loss-tolerant quantum weak coin-flipping protocols exist with arbitrarily small bias? This question will be addressed in a subsequent paper.

We have recently implemented our own loss-tolerant protocol and we have successfully tested it against Alice's and Bob's optimal cheating strategies. We report on those experiments elsewhere [32]. This was in a sense going full circle because it was our wish to implement Ambainis' quantum coin-flipping protocol that lead to the realization that we could never succeed and thus to the development of our protocol.

## ACKNOWLEDGMENTS

The authors are grateful to Hassène Bada for pointing out to us the fact that Spekkens and Rudolph had worked out a tight analysis of the ATVY protocol and to Serge Massar, Carlos Mochon, and Robert Spekkens for discussing their papers very patiently with us. Somshubhro Bandyopadhyay, Sébastien Gambs, and Tal Mor have helped the authors concerning the issue of maximum confidence quantum measurements. G.Br. was supported in part by Canada's Natural Sciences and Engineering Research Council (NSERC), the Canada Research Chair program, the Canadian Institute for Advanced Research (CIFAR), QuantumWorks and the Institut transdisciplinaire d'informatique quantique (INTRIQ). F.B. was supported in part by NSERC, the Fonds québécois de la recherche sur la nature et les technologies (FQRNT), and the Canadian Institute for Photonics Innovations (CIPI). N.G. was supported in part by the Centre d'optique, photonique et lasers (COPL), QuantumWorks, NSERC, CIPI, and INTRIQ.

- 
- [1] M. Blum, *Advances in Cryptology: A Report on CRYPTO'81* (Santa Barbara, California, 1981), pp. 11–15.
  - [2] P. W. Shor, *SIAM J. Comput.* **26**, 1484 (1997).
  - [3] A. Kent, *Phys. Rev. Lett.* **83**, 5382 (1999).
  - [4] C. H. Bennett and G. Brassard, *Proceedings of International Conference on Computers, Systems, and Signal Processing*, Bangalore, India, 1984, pp. 175–179.
  - [5] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
  - [6] H.-K. Lo and H. F. Chau, *Physica D* **120**, 177 (1998).
  - [7] D. Mayers, L. Salvail, and Y. Chiba-Kohno, e-print arXiv:quant-ph/9904078.
  - [8] D. Aharonov, A. Ta-Shma, U. Vazirani, and A. C.-C. Yao, *Proceedings of 32nd Annual ACM Symposium on Theory of Computing*, Portland, OR, 2000, pp. 705–714.
  - [9] R. W. Spekkens and T. Rudolph, *Quantum Inf. Comput.* **2**, 66 (2001).
  - [10] R. W. Spekkens and T. Rudolph, *Phys. Rev. A* **65**, 012310 (2001).
  - [11] A. Ambainis, *J. Comput. Syst. Sci.* **68**, 398 (2004).
  - [12] A. Kitaev, Lecture delivered at the 2003 Annual Quantum Information Processing (QIP) Workshop, Mathematical Sciences Research Institute, Berkeley, CA (unpublished). Available online at <http://www.msri.org/publications/video/index05.html>
  - [13] A. Chailloux and I. Kerenidis, *Proceedings of the 50th Annual IEEE Symposium on the Foundations of Computer Science*, Atlanta, GA, 2009.
  - [14] J. Barrett and S. Massar, *Phys. Rev. A* **69**, 022322 (2004).
  - [15] H. Buhrman, M. Christandl, M. Koucký, Z. Lotker, B. Patt-Shamir, and N. K. Vereshchagin, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, Lecture Notes in Computer Science, Vol. 4627 (Springer, Berlin, 2007), pp. 366–379.
  - [16] A. T. Nguyen, J. Frison, K. Phan Huy, and S. Massar, *New J. Phys.* **10**, 083037 (2008).
  - [17] D. Mayers, *Phys. Rev. Lett.* **78**, 3414 (1997).
  - [18] H.-K. Lo and H. F. Chau, *Phys. Rev. Lett.* **78**, 3410 (1997).
  - [19] E. Schrödinger, *Proc. Cambridge Philos. Soc.* **32**, 446 (1936).
  - [20] L. P. Hughston, R. Jozsa, and W. K. Wootters, *Phys. Lett. A* **183**, 14 (1993).
  - [21] K. A. Kirkpatrick, *Found. Phys. Lett.* **19**, 95 (2006).

- [22] C. A. Fuchs and J. van de Graaf, *IEEE Trans. Inf. Theory* **45**, 1216 (1999).
- [23] G. Brassard, C. Crépeau, R. Jozsa, and D. Langlois, *Proceedings of 34th Annual IEEE Symposium on the Foundations of Computer Science*, Palo Alto, CA, 1993, pp. 362–371.
- [24] C. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [25] W. Y. Hwang, D. Ahn, and S. W. Hwang, *Phys. Rev. A* **64**, 064302 (2001).
- [26] D. Dieks, *Phys. Lett. A* **126**, 303 (1988).
- [27] I. D. Ivanovic, *Phys. Lett. A* **123**, 257 (1987).
- [28] A. Peres, *Phys. Lett. A* **128**, 19 (1988).
- [29] A. Chefles and S. M. Barnett, *J. Mod. Opt.* **45**, 1295 (1998).
- [30] S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, *Phys. Rev. Lett.* **96**, 070401 (2006).
- [31] Y. Zhao, C.-H. F. Fung, B. Qi, C. Chen, and H.-K. Lo, *Phys. Rev. A* **78**, 042333 (2008).
- [32] G. Berlín, G. Brassard, F. Bussiès, N. Godbout, J. A. Slater, and W. Tittel, e-print arXiv:0904.3946.
- [33] C. Mochon, e-print arXiv:0711.4114.