Security bound of cheat sensitive quantum bit commitment

Guang Ping He

School of Physics and Engineering, Sun Yat-sen University, Guangzhou 510275, China.

Quantum bit commitment (QBC) is a two-party cryptography including the following phases. In the commit phase, Alice (the sender of the commitment) decides the value of the bit \( b \) \((b = 0 \text{ or } 1)\) that she wants to commit, and sends Bob (the receiver of the commitment) a piece of evidence, e.g., some quantum states. Later, in the unvail phase, Alice announces the value of \( b \), and Bob checks it with the evidence. The interval between the commit and unveil phases is sometimes called the holding phase. A QBC protocol is called unconditionally secure if any cheating can be detected with a probability arbitrarily close to 1. Here Alice’s cheating means that she wants to change the value of \( b \) after the commit phase, while Bob’s cheating means that he tries to learn \( b \) before the unveil phase.

QBC is an essential primitive for building quantum multi-party secure computations and other “post-cold-war era” multi-party cryptographic protocols. Unfortunately, it is widely believed that unconditionally secure QBC is impossible. This result, known as the Mayers-Lo-Chau (MLC) no-go theorem, was considered as putting a serious drawback on quantum cryptography.

To evade the problem, the concept “cheat sensitive quantum bit commitment (CSQBC)” was proposed, where the probability for detecting the cheating does not need to be arbitrarily close to 1. Instead, it merely requires the probability to be nonzero. With this loosen security requirement, many insecure QBC protocols can be regarded as secure CSQBC. Therefore, at the first glance it seems that CSQBC will be very easy to achieve. But intriguingly, here we will show that there still exists boundary for the security of a typical class of CSQBC. Especially, Bob can always feel free to measure the quantum states to learn \( b \), while he stands at least 50% chances to escape Alice’s detection.

Results

**Common features of CSQBC.** By checking the existing CSQBC protocols, we find that they all share the following common features (note that the names Alice and Bob are used reversely in Refs. 7, 9, 10):

1. During the holding phase, the receiver Bob owns a quantum system \( \Psi \) encoding Alice’s committed bit \( b \). (\( \Psi \) can either be prepared by the sender Alice, or be prepared by Bob and sent to Alice, who returns it to Bob after performing some certain operations according to her choice of \( b \). It also does not matter whether Alice prepared and kept another quantum system entangling with \( \Psi \).)

2. Bob knows the definitions of \( \rho^0 \) and \( \rho^1 \) directly before the end of the commit phase. (That is, these definitions are either clearly stated by the protocol, or announced to Bob by Alice classically. Bob does not need to perform operations on any quantum system to gain knowledge of these definitions.) Here \( \rho^0 \) and \( \rho^1 \) are the density matrices of Bob’s \( \Psi \) corresponding to \( b = 0 \) and \( b = 1 \), respectively.

3. To detect Bob’s cheating, at the unveil phase Alice can check whether the state of \( \Psi \) is intact. (It does not matter whether the entire \( \Psi \) or only a small part can be checked.)

4. To detect Alice’s cheating, at the unveil phase Bob can learn a nontrivial amount of information on the value of \( b \) from \( \Psi \), even without any help from Alice.
The last feature indicates that there exists at least one operation known to Bob, which can output a bit \( b' \) when being applied on \( \Psi \), and \( b' = b \) should occur with a probability larger than 1/2. As a result, there must be \( \rho_0^b \neq \rho_1^b \). This is a major difference from the original QBC, where there is generally \( \rho_0^b \approx \rho_1^b \) so that it can be unconditionally secure against dishonest-Bob.

The original purpose of CSQBC having these features is as follows. Alice’s cheating strategy suggested in the MLC no-go theorem is based on the Hughston-Jozsa-Wootters (HJW) theorem\(^{11} \), which applies to the case \( \rho_0^b \approx \rho_1^b \). Therefore, with feature (4), i.e., \( \rho_0^b \neq \rho_1^b \), Alice’s cheating becomes detectable so that the MLC no-go theorem can be evaded. On the other hand, if Bob takes advantages of \( \rho_0^b \neq \rho_1^b \) and performs measurements to discriminate the committed bit \( b \), the quantum state will be disturbed. In this case, with feature (3) Bob’s cheating will be detected with a certain probability when Alice asks him to return the quantum state and checks whether it remains undisturbed, so that the goal of CSQBC can be met.

But with a rigorous quantitative analysis on the probability of detecting Bob’s cheating, we will find that it is always not sufficiently large when Bob applies some specific measurements. Therefore any CSQBC protocol having the above four features will be bounded by the security limit below.

**Notations and Bob’s cheating strategy.** According to Eq. (9.22) of Ref. 12, the trace distance \( D(\rho_0^b, \rho_1^b) = \| \rho_0^b - \rho_1^b \| / \sqrt{2} \) (where \( |A| = \sqrt{\text{tr} A^2} \) between \( \rho_0^b \) and \( \rho_1^b \) satisfies

\[
D(\rho_0^b, \rho_1^b) = \max \text{tr} \left( \rho_0^b - \rho_1^b \right),
\]

where the maximization is taken over all positive operators \( P \leq I \), with \( I \) being the identity operator. The above feature (2) of CSQBC guarantees that Bob knows how \( \rho_0^b \) and \( \rho_1^b \) are defined. Thus he can find the positive projectors \( P = P_m \) that maximizes \( \text{tr} \left( \rho_0^b - \rho_1^b \right) \). If \( \rho_0^b \) stands a higher probability to be projected successfully than \( \rho_1^b \) when applying \( P_m \), then we take \( P_0 = P_m \) and \( P_1 = I - P_m \). Otherwise we take \( P_0 = I - P_m \) and \( P_1 = P_m \). Feature (1) ensures that Bob owns the system \( \Psi \) encoding Alice’s committed bit \( b \) during the holding phase. Therefore, by applying the positive operator-valued measure (POVM) \( \left\{ P_0, P_1 \right\} \) on \( \Psi \), Bob can discriminate between \( \rho_0^b \) and \( \rho_1^b \) and learn Alice’s committed bit \( b \) with the maximum probability allowed by \( D(\rho_0^b, \rho_1^b) \).

To analyze rigorously the probability for Bob to escape Alice’s detection with this POVM, let \( H \) be the global Hilbert space constructed by all possible states of \( \Psi \) (either \( b = 0 \) or \( 1 \)). Since \( P_0, P_1 \) are positive projectors, there exists an orthonormal basis \( \{ |i\rangle \} \) of \( H \) (the following proof remains valid regardless whether \( |\eta\rangle \) is known to Alice or Bob), in which \( P_0, P_1 \) can be expressed as

\[
P_0 = \sum_i |i^{(0)}\rangle \langle i^{(0)} |,
\]

\[
P_1 = \sum_i |i^{(1)}\rangle \langle i^{(1)} |,
\]

where \( \{ |i^{(0)}\rangle \} \cup \{ |i^{(1)}\rangle \} = \{ |i\rangle \} \).

Meanwhile, before Bob applying any measurement, the general form of the initial state of \( \Psi \) can always be written as

\[
|\Phi \otimes \Psi\rangle_{im} = \sqrt{\alpha} \sum_i |i^{(0)}\rangle |f_i^{(0)}\rangle \otimes |e_i^{(0)}\rangle + \sqrt{\beta} \sum_i |i^{(1)}\rangle |f_i^{(1)}\rangle \otimes |e_i^{(1)}\rangle,
\]

where \( 0 \leq \alpha \leq 1, \beta = 1 - \alpha, \) and \( \sum_i |e_i^{(0)}\rangle^2 = \sum_i |e_i^{(1)}\rangle^2 = 1 \) (sum over all possible \( i \) within each corresponding subspace). The values of \( \alpha, \beta, |f_i^{(0)}\rangle \)s and \( |f_i^{(1)}\rangle \)s are chosen by Alice according to the value of her committed bit \( b \). Here \( \Phi \) is a quantum system that Alice may introduce and keep to herself, which entangles with Bob’s \( \Psi \). All \( |f_i^{(0)}\rangle \)s and \( |f_i^{(1)}\rangle \)s are the vectors describing the state of \( \Phi \), which are not required to be orthogonal to each other. In the case where Alice does not introduce such a system, we can simply set all \( |f_i^{(0)}\rangle \)s and \( |f_i^{(1)}\rangle \)s to be equal, so that Eq. (3) still applies.

The security bound on Bob’s cheating. As elaborated in the 1st subsection of Methods section, when dishonest-Bob applies the above POVM \( \{ P_0, P_1 \} \) on \( \Psi \), we find that the probability for Bob’s cheating to pass Alice’s detection successfully is

\[
P_b = \frac{1}{2} + \frac{1}{2} (2x - 1)^2,
\]

and the amount of mutual information he obtained is

\[
I_m = 1 - h(x).
\]

Here \( h(x) = -x \log_2 x - (1 - x) \log_2 (1 - x) \) is the binary entropy function.

With Eqs. (4) and (5), we plot \( P_b \) and \( I_m \) as a function of \( x \) in FIG. 1. Since \( 0 \leq x \leq 1 \), FIG. 1 and Eq. (4) both give

\[
P_b \geq 50%.
\]

The minimum \( P_b = 50% \) can be reached when Alice chooses \( x = 0.5 \). Thus we come to the conclusion that Bob can always learn Alice’s committed bit \( b \) with the maximal probability allowed by the trace distance between \( \rho_0^b \) and \( \rho_1^b \), while his cheating stands at least 50% chance to escape Alice’s detection.

It may look weird that FIG. 1 seems to indicate that the more information that Bob obtains, the easier he can pass Alice’s detection. But we must note that the amount of Bob’s information is not chosen by himself. Instead, it is determined by the value of \( x \) that Alice chooses. That is, once Alice determines which state is used for encoding her committed bit, the maximum amount of information that Bob can obtain is also fixed.

On the other hand, the above result indicates that Alice should make \( x \) as close to 0.5 as possible, so that Bob’s information and successful cheating probability can be minimized. However, note that she has to choose the initial state Eq. (3) within the range restricted by the protocol. Due to the feature (4) of CSQBC, the trace distance \( D(\rho_0^b, \rho_1^b) \) has to be nonzero. Therefore, generally \( x \) cannot be made very close to 0.5, as we will see in the examples below.

**Examples.** In the CSQBC protocol in Ref. 5, Bob’s system \( \Psi \) is a single qubit, whose state is either \( |0\rangle \) or \( \pm |1\rangle \) when Alice commits \( b = 0 \) (\( b = 1 \)). Here \( |0\rangle \) and \( |1\rangle \) are orthogonal to each other, \( |\pm\rangle \equiv (|0\rangle \pm |1\rangle) / \sqrt{2} \). So we have \( \rho_0^b = (|0\rangle \langle 0| + |1\rangle \langle -1|) / 2\) and \( \rho_1^b = (|1\rangle \langle 1| + + |1\rangle \langle +|) / 2\). Define

\[
|e^{(0)}\rangle \equiv \cos(\pi/8) |0\rangle + \sin(\pi/8) |1\rangle,
\]

\[
|e^{(1)}\rangle \equiv \cos(3\pi/8) |0\rangle + \sin(3\pi/8) |1\rangle.
\]

Then Bob’s operation for maximally discriminating \( \rho_0^b \) and \( \rho_1^b \) is to measure \( \Psi \) in the basis \( \{ |e^{(0)}\rangle, |e^{(1)}\rangle \} \), i.e., he applies the projector \( P_0 = |e^{(0)}\rangle \langle e^{(0)}| \). When the projection is successful (unsuccessful), he takes \( b' = 0 \) (\( b' = 1 \)) as the decoded result. With this method,
b’ will match Alice’s actual committed bit b with the probability \( \cos^2(\pi/8) \approx 85.36\% \). Meanwhile, Alice’s four input states can be expanded in the \(|e^{0}\rangle, |e^{1}\rangle\) basis as

\[
\begin{align*}
|0\rangle &= \cos(\pi/8)|e^{0}\rangle + \sin(\pi/8)|e^{1}\rangle, \\
|\rangle &= \cos(\pi/8)|e^{0}\rangle - \sin(\pi/8)|e^{1}\rangle, \\
|1\rangle &= -\sin(\pi/8)|e^{0}\rangle + \cos(\pi/8)|e^{1}\rangle, \\
|+\rangle &= \sin(\pi/8)|e^{0}\rangle + \cos(\pi/8)|e^{1}\rangle.
\end{align*}
\]

Comparing with Eq. (3), we can see that there is either \( \alpha = \cos^2(\pi/8) \) or \( \alpha = \sin^2(\pi/8) \). Substitute them into Eq. (4) will both yield \( P_b = 85.36\% \) in Ref. 5. That is, in the CSQBC protocol, Bob’s system \( \Psi \), there is another system \( E \). Alice’s different committed values of \( b \) is encoded with different states of the combined system \( E \otimes \Psi \). System \( E \) is kept at Alice’s side during the commit and holding phases, and is required to be sent to Bob at the unveil phase to justify Alice’s commitment. Let \( \rho_{0}^{EB} \) and \( \rho_{1}^{EB} \) denote the density matrices of \( E \otimes \Psi \) corresponding to \( b = 0 \) and \( b = 1 \), respectively. Note that in all existing CSQBC protocols, there is no such a system \( E \). But we include it here, so that the model can cover more protocols that may be proposed in the future.

In this scenario, Alice’s cheating strategy is as follows. At the beginning of the protocol she introduces an ancillary system \( \Phi \) which is a copy of \( E \otimes \Psi \). Since the fidelity

\[
F(\rho_{0}^{EB}, \rho_{1}^{EB}) \equiv tr\sqrt{(\rho_{0}^{EB})^{1/2}(\rho_{1}^{EB})^{1/2}}
\]

between \( \rho_{0}^{EB} \) and \( \rho_{1}^{EB} \) satisfies

\[
F(\rho_{0}^{EB}, \rho_{1}^{EB}) = \max_{\langle \psi_{0}\rangle, \langle \psi_{1}\rangle} |\langle \psi_{0}|\psi_{1}\rangle|,
\]

where the maximization is over all purifications \( |\psi_{0}\rangle \) of \( \rho_{0}^{EB} \) and \( |\psi_{1}\rangle \) of \( \rho_{1}^{EB} \) into \( \Phi \otimes E \otimes \Psi \), Alice finds the real and positive \( |\psi_{0}\rangle, |\psi_{1}\rangle \) that reach the maximum, i.e.,

\[
F(\rho_{0}^{EB}, \rho_{1}^{EB}) = \langle \psi_{0}|\psi_{1}\rangle = |\langle \psi_{0}|\psi_{1}\rangle|
\]

Then she prepares the initial state of \( \Phi \otimes E \otimes \Psi \) as

\[
|\psi_{c}\rangle = \frac{|\psi_{0}\rangle + |\psi_{1}\rangle}{N}
\]

where the normalization constant

\[
N = \sqrt{2 + \langle \psi_{0}|\psi_{1}\rangle + \langle \psi_{1}|\psi_{0}\rangle}.
\]

She uses this state to complete the rest of the commit protocol. With this method, the value of \( b \) is not determined during the commit phase.

In the unveil phase, Alice decides whether she wants to unveil \( b = 0 \) or \( b = 1 \). Then she simply uses \( |\psi_{c}\rangle \) as \( |\psi_{b}\rangle \) to complete the

Figure 1 | Bob’s successful cheating probability \( P_b \) (red line) and mutual information \( I_m \) (blue line) on Alice’s committed bit \( b \) as a function of \( \alpha \) that Alice chooses for the initial state Eq.(3). The dash lines mark the values for the protocol in Ref. 5.

Alice’s cheating strategy. Alice’s cheating strategy used in the MLC no-go theorem requires the condition \( \rho_{0}^{EB} \approx \rho_{1}^{EB} \), which no longer holds in CSQBC. Nevertheless, she can still apply the same strategy in CSQBC and try her luck. To give a detailed description of the strategy, first we model the coding method in CSQBC more precisely. For generality, consider that in the protocol, besides Bob’s system \( \Psi \), there is another system \( E \). Alice’s different committed values of \( b \) is encoded with different states of the combined system \( E \otimes \Psi \). System \( E \) is kept at Alice’s side during the commit and holding phases, and is required to be sent to Bob at the unveil phase to justify Alice’s commitment. Let \( \rho_{0}^{EB} \) and \( \rho_{1}^{EB} \) denote the density matrices of \( E \otimes \Psi \) corresponding to \( b = 0 \) and \( b = 1 \), respectively. Note that in all existing CSQBC protocols, there is no such a system \( E \). But we include it here, so that the model can cover more protocols that may be proposed in the future.

In this scenario, Alice’s cheating strategy is as follows. At the beginning of the protocol she introduces an ancillary system \( \Phi \) which is a copy of \( E \otimes \Psi \). Since the fidelity

\[
F(\rho_{0}^{EB}, \rho_{1}^{EB}) \equiv tr\sqrt{(\rho_{0}^{EB})^{1/2}(\rho_{1}^{EB})^{1/2}}
\]

between \( \rho_{0}^{EB} \) and \( \rho_{1}^{EB} \) satisfies

\[
F(\rho_{0}^{EB}, \rho_{1}^{EB}) = \max_{\langle \psi_{0}\rangle, \langle \psi_{1}\rangle} |\langle \psi_{0}|\psi_{1}\rangle|,
\]

where the maximization is over all purifications \( |\psi_{0}\rangle \) of \( \rho_{0}^{EB} \) and \( |\psi_{1}\rangle \) of \( \rho_{1}^{EB} \) into \( \Phi \otimes E \otimes \Psi \), Alice finds the real and positive \( |\psi_{0}\rangle, |\psi_{1}\rangle \) that reach the maximum, i.e.,

\[
F(\rho_{0}^{EB}, \rho_{1}^{EB}) = \langle \psi_{0}|\psi_{1}\rangle = |\langle \psi_{0}|\psi_{1}\rangle|
\]

Then she prepares the initial state of \( \Phi \otimes E \otimes \Psi \) as

\[
|\psi_{c}\rangle = \frac{|\psi_{0}\rangle + |\psi_{1}\rangle}{N}
\]

where the normalization constant

\[
N = \sqrt{2 + \langle \psi_{0}|\psi_{1}\rangle + \langle \psi_{1}|\psi_{0}\rangle}.
\]

She uses this state to complete the rest of the commit protocol. With this method, the value of \( b \) is not determined during the commit phase.

In the unveil phase, Alice decides whether she wants to unveil \( b = 0 \) or \( b = 1 \). Then she simply uses \( |\psi_{c}\rangle \) as \( |\psi_{b}\rangle \) to complete the
Figure 2 | The lower bound of the sum of the cheating probabilities \( P_A + P_B \) as a function of the trace distance \( D(\rho^B_0, \rho^B_1) \) and the probability \( \zeta \) with which Bob’s action is checked. \( A \) and \( B \) denote the points that reach the minimum \( P_A + P_B = 3/2 \).

Protocol. From the symmetry of \( |\psi_0 \rangle \) and \( |\psi_1 \rangle \) in Eq.(11), we can see that her successful cheating probabilities for \( b = 0 \) and \( b = 1 \) are both

\[
P_A = |\langle \psi_0 | \psi_1 \rangle|^2 = \frac{1 + |\langle \psi_0 | \psi_1 \rangle(1 + |\langle \psi_1 | \psi_0 \rangle)|}{2 + |\langle \psi_0 | \psi_1 \rangle + |\langle \psi_1 | \psi_0 \rangle|} \tag{13}
\]

Therefore, in any specific CSQBC protocol, the Alice’s exact cheating probability can be calculated once the definition of \( \rho^B_0, \rho^B_1 \) is known.

The optimal protocols are trivial. Now we will try to find the CSQBC protocols which can optimally detect the cheating of both parties, i.e., minimizing the sum of Alice’s and Bob’s cheating probabilities.

Note that Eq. (4) depends on the specific value of \( z \) in the state Eq. (3) that Alice chooses in a single run of the protocol, while \( F(\rho^B_0, \rho^B_1) \) in Eq. (13) is the statistical result of all the legitimate states allowed by the protocol. Thus it is hard to compare Eq. (13) and Eq. (4) directly and give a general result without knowing the details on the composition of \( \rho^B_0 \) in a specific protocol.

Fortunately, in all existing CSQBC protocols\(^6–10\), there is no system \( E \). The form of the states of Bob’s system \( \Psi \) alone carries all the information of \( b \). Thus the trace distance \( D(\rho^E_0, \rho^E_1) = D(\rho^B_0, \rho^B_1) \). For any protocol of this kind (as well as protocols having system \( E \) but still satisfying \( D(\rho^E_0, \rho^E_1) = D(\rho^B_0, \rho^B_1) \)), we can replace both \( z \) and \( F(\rho^B_0, \rho^B_1) \) with \( D(\rho^B_0, \rho^B_1) \), as elaborated in the 2nd subsection of Method, where we obtain

\[
P_A \geq \frac{1 - D(\rho^B_0, \rho^B_1)}{2}, \tag{14}
\]

and

\[
P_B \geq \frac{1 + D(\rho^B_0, \rho^B_1)}{2}. \tag{15}
\]

These two equations suggest that \( P_A \) and \( P_B \) cannot be minimized simultaneously in the same protocol, because reducing \( P_A \) requires a higher \( D(\rho^B_0, \rho^B_1) \), while it will result in a higher \( P_B \) at the same time.

Moreover, we must note that the above \( P_A \) and \( P_B \) are obtained assuming that the actions of both parties in the protocol will always be checked. But this is impossible, because they share the same system \( \Phi \otimes E \otimes \Psi \). In the unveil phase, either Bob will measure \( E \otimes \Psi \) to check Alice’s action, or he is required to return \( \Psi \) to Alice who checks his action. These cannot be done simultaneously. Suppose that in a CSQBC protocol, Bob’s action is checked with probability \( \zeta (0 \leq \zeta \leq 1) \), and Alice’s action is checked with probability \( 1 - \zeta \). When one’s action is not checked, he/she can cheat successfully with probability 1. Thus the cheating probabilities \( P_A \) and \( P_B \) should be replaced by

\[
P_A' = \zeta + (1 - \zeta) P_A \tag{16}
\]

and

\[
P_B' = (1 - \zeta) + \zeta P_B, \tag{17}
\]

respectively. Combining them with Eqs. (14) and (15), we find

\[
P_A' + P_B' \geq \frac{\zeta + D(\rho^B_0, \rho^B_1)}{2} + \frac{1 + D(\rho^B_0, \rho^B_1)}{2} \tag{18}
\]

Since \( 0 \leq \zeta \leq 1 \) and \( 0 \leq D(\rho^B_0, \rho^B_1) \leq 1 \), we find another security lower bound of CSQBC

\[
P_A' + P_B' \geq \frac{3}{2}. \tag{19}
\]

To find the optimal protocol that can reach this bound, we plot the lower bound of \( P_A' + P_B' \) as a function of \( D(\rho^B_0, \rho^B_1) \) and \( \zeta \) in Fig. 2 according to Eq. (18). It shows that there are two types of protocols that can both reach the minimum \( P_A' + P_B' = 3/2 \), denoted as points \( A \) and \( B \) in Fig. 2, respectively, with the parameters (A) \( D(\rho^B_0, \rho^B_1) = 1, \zeta = 0 \), and (B) \( D(\rho^B_0, \rho^B_1) = 0, \zeta = 1 \). Type (A) protocols mean that \( \rho^B_0 \) and \( \rho^B_1 \) are orthogonal so that \( P_A \) reaches its minimum 1/2. However, \( \rho^B_0 \) and \( \rho^B_1 \) can be distinguished perfectly and Bob’s action is never checked. Thus \( P_B = 1 \), i.e., he can always learn Alice’s committed \( b \) with reliability 1 and never get caught. In type (B) protocols, \( \rho^B_0 = \rho^B_1 \), so that Bob learns nothing about \( b \). But Alice’s action is never checked so that she can unveil \( b \) as whatever she wants, with a successful cheating probability \( P_A' = 1 \). Therefore, we can see that these optimal protocols are all trivial as they are.
completely insecure against one of the parties. Thus they do not seem to have any practical usage.

The fair protocol. Since the protocols that can minimize $P_A^* + P_B^*$ all look useless, let us consider the protocol satisfying $P_A^* = P_B^*$ so that it is fair for both parties, and try to minimize $P_A^*, P_B^*$ in this case. From Eq. (37) we can see that the inequality Eq. (15) can become equality when $\alpha^2 = \alpha$, i.e., all the states allowed to be chosen in the protocol for committing the same $b$ value should have the same $\alpha$ value. Also, note that the lowest bounds in Eqs. (14) and (18) cannot be reached by most $D(\rho_0^B, \rho_1^B)$, because these inequalities can become equalities if and only if $F(\rho_0^B, \rho_1^B) = 1 - D(\rho_0^B, \rho_1^B)$, which requires $\rho_0^B = \rho_1^B$. Therefore, only the above optimal protocols can reach these bound. For this reason, to calculate $P_A^*$ in other protocols, we should use Eq. (13) instead of Eq. (14). To compute $F(\rho_0^B, \rho_1^B)$ in Eq. (13), for simplicity we consider only the protocols in which there are

$$\rho_0^B = \rho_1^B = 0 \begin{bmatrix} \alpha & 0 \\ 0 & 1 - \alpha \end{bmatrix},$$

$$\rho_1^B = 1 - \alpha \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

then

$$F(\rho_0^B, \rho_1^B) = 2\sqrt{\alpha(1 - \alpha)},$$

$$D(\rho_0^B, \rho_1^B) = 2\alpha - 1.$$ (21)

Combining them with Eqs. (13), (16), (17) and (15) (the latter becomes equality once we choose $\alpha^2 = \alpha$), then by solving $P_A^* = P_B^*$ we yield

$$\tilde{\gamma} = \frac{2\sqrt{\alpha(1 - \alpha) - 1}}{(2\alpha - 1)^2 + 2\sqrt{\alpha(1 - \alpha) - 2}}.$$ (22)

Any protocol satisfying this equation is fair for both parties. Now let us find the minimal value of $P_A^* = P_B^*$. Substituting this $\tilde{\gamma}$ into Eq. (17), we obtain

$$P_A^* = P_B^* = \frac{(2\sqrt{\alpha(1 - \alpha) + 1})(2\alpha^2 - 2\alpha + 1) - 2}{4\alpha^2 - 2\alpha + 2\sqrt{\alpha(1 - \alpha) - 1}}.$$ (23)

By solving $dP_A^*/d\alpha = 0$, we find that the minimal cheating probabilities in such protocols are $P_A^* = P_B^* = 0.904$, which can be obtained when $\alpha \approx 0.885$, i.e., $\sqrt{\alpha} \approx 0.941$ of $19.85\%$. In this case $\tilde{\gamma} \approx 0.469$.

A simple protocol having these parameters is: Alice sends Bob the state $|0\rangle \pm \sin(19.85\%) |1\rangle$ if she wants to commit $b = 0$ ($b = 1$). In the unveil phase, with probability $\tilde{\gamma} \approx 0.469$ Bob returns the state and Alice checks whether it remains undisturbed, with probability $1 - \tilde{\gamma} \approx 0.531$ Bob measures the state and checks whether it agrees with the value of Alice unveiled.

Nevertheless, there is the difficulty in finding a method for deciding which party will be checked in a single run of the protocol. Dishonest Alice (Bob) would like to decrease $1 - \tilde{\gamma}$ (\tilde{\gamma}) so that $P_A^* (P_B^*)$ can be raised. Thus they do not trust each other and may not collaborate. The CSQBC protocol in Ref. 5 adopts a process called “the game” to handle this problem, which is very similar to quantum coin flipping (QCF) protocols\(^1\). However, Ishizaka\(^1\) showed that this process provides extra security loophole to Bob, so that there is a cheating strategy for him to learn $b$ with reliability 61.79% (which is lower than what can be obtained with our cheating strategy, as calculated in the Examples section) while passing Alice’s check with probability 100% (which is higher than that of our strategy). It was further shown in Ref. 16 that due to the existence of ideal black-boxed QCF, any CSQBC protocol based on biased QCF cannot be secure. Therefore, it remains unclear how to build a fair CSQBC protocol with $P_A^* = P_B^*$ while minimizing $P_A^*$ and $P_B^*$.

Discussion

In summary, we showed that any CSQBC protocol having the above four features is subjected to the security bound Eq. (6). Protocols satisfying $D(\rho_0^B, \rho_1^B) = D(\rho_0^B, \rho_1^B)$ is further bounded by Eq. (19). Note that the insecurity of QCF-based CSQBC protocols (e.g., Refs. 5, 6) was already pinpointed in Refs. 15, 16. But our proof also applies to the non-QCF-based ones.

Our result should not be simply considered as a generalization of the MLC no-go proof. Instead, it is a complement. This is because the MLC no-go proof applies to QBC protocol with $P_A^* \approx P_B^*$, but as pointed out in Ref. 9, CSQBC does not need to satisfy this requirement so that it may evade the MLC theorem. On the contrary, our proof works for the case $P_A^* \neq P_B^*$, thus it fills the gap where the MLC proof left. Meanwhile, the MLC theorem concentrates on the cheating of Alice. It does not exclude the existence of protocols which is unconditionally secure against dishonest Bob only. On the other hand, our result shows that Bob can always cheat in CSQBC regardless Alice is honest or not.

It will be interesting to study whether there can be CSQBC protocols without the above four features. It seems that Kent’s relativistic QBC\(^17\)-\(^19\) and our recent proposals\(^20\)-\(^23\) do not satisfy feature (1), while the protocol in Ref. 22 does not have feature (2), as elaborated in Ref. 23. However, these works are aimed to achieve the original QBC, instead of CSQBC. Also, Refs. 20–23 have not gained wide recognition yet. Thus it is still an open question whether it is possible to build non-relativistic CSQBC protocols which are not limited by the above security bounds, without relying on computational and experimental constraints.

Methods

Calculating Bob’s cheating probability. Consider the POVM $\{P_0, P_1, P_2, P_3\}$ defined in Eq. (2). After Bob applies it on $\Psi$, there can be two outcomes.

(I) The projection outcome is $P_0$. Then Bob takes $b' = 0$ as his decoded result of Alice’s committed bit $b$. With Eqs. (2) and (3) we yield

$$P_b = \sum_i |i\rangle \langle i| \otimes |i\rangle \langle i|.$$ (24)

Thus this case will occur with the probability

$$p_b = \frac{|1}{\sum_i |i\rangle \langle i|}.$$ (25)

while the resultant state of $\Phi \otimes \Psi$ is

$$|\Phi \otimes \Psi\rangle = \frac{1}{p_b} P_b |\Phi \otimes \Psi\rangle.$$ (26)

As described in feature (3) of CSQBC, at the unveiling phase Alice may require Bob to return $\Psi$ and check whether it remains intact in its initial state. The maximal probability for Alice to find out that Bob has already projected $|\Phi \otimes \Psi\rangle$ into $|\Phi \otimes \Psi\rangle$ is bounded by

$$P_{\pi} = 1 - \frac{1}{p_b}.$$ (27)

Thus the total probability for (case (I) occurred) AND (Alice failed to detect Bob’s cheating) is

$$p_t = (1 - P_{\pi}).$$ (28)

(II) The projection outcome is $P_1$. Then Bob takes $b' = 1$ as his decoded result of Alice’s $b$. Now

$$P_{\pi} = P_{\pi} = \sqrt{\sum_i |j\rangle \langle j|}.$$ (29)
Obviously, this case will occur with the probability
\[ p_1 = 1 - p_0. \] (30)

Meanwhile, the resultant state of $\Phi \otimes \Psi$ in this case is
\[ |\Phi \otimes \Psi\rangle_{ii} = \frac{1}{\sqrt{p_0}} P_1 |\Phi \otimes \Psi\rangle_{ss}. \] (31)

The maximal probability for Alice to find out that Bob has already projected $|\Phi \otimes \Psi\rangle_{ss}$ into $|\Phi \otimes \Psi\rangle_{ii}$ is bounded by
\[ \bar{p}_B = 1 - |\langle \Phi \otimes \Psi | (\Phi \otimes \Psi)_{ss} \rangle^2 | \]
\[ = 1 - \frac{1}{p_0} \beta^2, \] (32)

Thus the total probability for (case (II) occurred) AND (Alice failed to detect Bob’s cheating) is
\[ p_1 (1 - \bar{p}_B) = \beta^2. \] (33)

Taking both cases (I) and (II) into consideration, the overall probability for Bob’s cheating to pass Alice’s detection successfully is
\[ p_B = p_1 (1 - \bar{p}_B) + p_0 (1 - \bar{p}_0) = x^2 + \beta^2 \]
\[ = 1 + \frac{1}{2} (2x - 1)^2. \] (34)

Meanwhile, since the projection outcome will either be $p_0$ or $p_1$ with the probabilities $p_1$ and $p_0 = 1 - p_1$, respectively, Bob’s $b’$ will match Alice’s $b$ with the probability $p_1$ or $1 - p_1$ too. Note that $h(1 - p_0) = h(p_0)$. Thus the amount of mutual information that Bob obtains with this POVM is
\[ I_n = 1 - h(p_1) = 1 - h(x). \] (35)

Bounding the cheating probabilities with trace distance. Suppose that there are many states allowed to be chosen randomly for committing $b = 0$ in the protocol, each of which takes the form of Eq. (3), but with different values of the coefficients $\beta, \lambda^{(b_0)}_j, \lambda^{(b_1)}_j$. Bob applies the optimal POVM to decode $b$. Then Eq. (3) indicates that he can learn $b$ correctly with probability $\bar{p}$, i.e., the average of $x$. Meanwhile, it is well-known that the maximal probability for discriminating two density matrices $\rho_B$ and $\rho'_B$ is $(1 + D(\rho_B, \rho'_B)) / 2$. Therefore
\[ D(\rho_B, \rho'_B) = 2x - 1. \] (36)

Since Eq. (4) shows that Bob’s average cheating probability for these states is
\[ p_B = \frac{1 + (2x - 1)^2}{2} \geq \frac{1 + (2x - 1)^2}{2}, \] (37)

we have
\[ p_B \geq 1 + D(\rho_B, \rho'_B)^2 \]
\[ \geq 2. \] (38)

Similar discussion is also valid for the states for committing $b = 1$, except that $\bar{z}$ should be replace by $\beta = 1 - \bar{z}$. But Eq. (38) remains the same because Eq. (4) satisfies $P_b(1 - \bar{z}) = P_b(z)$.

On the other hand, since
\[ F(\rho_B, \rho'_B) \geq 1 - D(\rho_B, \rho'_B) \]
from Eq. (13) we yield
\[ p_A \geq 1 - D(\rho_B, \rho'_B) \]
(40)