Security proof of counterfactual quantum cryptography against general intercept-resend attacks and its vulnerability*

Zhang Sheng(张 盛)†, Wang Jian(王 剑), and Tang Chao-Jing(唐朝京)

School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073*, China*

(Received 23 October 2011; revised manuscript received 4 January 2012)

Counterfactual quantum cryptography, recently proposed by Noh, is featured with no transmission of signal particles. This exhibits evident security advantages, such as its immunity to the well-known photon-number-splitting attack. In this paper, the theoretical security of counterfactual quantum cryptography protocol against the general intercept-resend attacks is proved by bounding the information of an eavesdropper Eve more tightly than in Yin's proposal [*Phys. Rev.* A **82** 042335 (2010)]. It is also shown that practical counterfactual quantum cryptography implementations may be vulnerable when equipped with imperfect apparatuses, by proving that a negative key rate can be achieved when Eve launches a time-shift attack based on imperfect detector efficiency.

Keywords: quantum cryptography, quantum counterfactuality, quantum information

PACS: 03.67.Dd, 03.67.Hk, 42.50.Ex **DOI:** 10.1088/1674-1056/21/6/060303

1. Introduction

Quantum key distribution (QKD),[1−3] which is the most prominent application in quantum information theory, enables two distant parties, conventionally referred to Alice and Bob, to establish a secret key guaranteed by fundamental quantum mechanics, such as the no-cloning theorem. Since the first QKD scheme was proposed in 1984,[1] a lot of attention has been payed both in theoretical and experimental areas in the following decades.[4−10] In most of the schemes, the information bits are conventionally encoded into the quantum states, e.g., chosen from two conjugated bases and then transmitted to Bob in the public channel controlled by a powerful eavesdropper Eve. Recently Noh proposed a novel QKD protocol (Noh09 protocol)[11] using a striking phenomenon commonly termed as the quantum counterfactual effect, which is initiated by the idea of interaction-free measurement in quantum computation.[12,13] It is exciting that the information particles in his protocol are not transmitted via any physical channel, thus, the security is witnessed in a straightforward manner.

In the Noh09 protocol, Alice prepares a single photon pulse from two orthogonal states, i.e., a horizontally polarized state $|H\rangle$ and a vertically polarized state $|V\rangle$. The pulse is then split into two by a beam splitter. The initial quantum states after the BS are written $|\phi_{0(1)}\rangle = \sqrt{T}|0\rangle_a|H(V)\rangle_b + \mathrm{i}\sqrt{R}|H(V)\rangle_a|0\rangle_b$, where $|0\rangle_{a(b)}$ denotes the vacuum state in mode $a(b)$. In Bob's secure zone, a polarization beam splitter (PBS) and optical switch (SW) are applied to block path b according to Bob's randomly chosen polarization. Specifically, the sub-pulse in path b is blocked only if their polarizations are identical, otherwise it will be reflected back to Alice. Ideally, there are three intrinsic events: (E_1) Detector D_1 clicks, occuring when the single photon is reflected by the BS and their polarizations are identical; (E_2) Detector D_2 clicks, occuring when Bob's polarization is inconsistent with Alice's, or when Bob's polarization is identical to Alice's and the photon travels along path a; (E_3) Detector D_3 clicks, occuring when Bob's polarization is identical to Alice's and the photon travels along path b. The raw key is generated from part of the instances in event E_1. Although a simple security analysis against a special intercept-resend attack was presented in the Noh09 protocol, a more strict one is expected.

Recently, Yin *et al.*[14] gave a security proof of the counterfactual quantum cryptography with the same technique proposed in Ref. [15]. In Yin's proposal,

*Project supported by the National Natural Science Foundation of China (Grant No 60872052).

†Corresponding author. E-mail: shengzhcn@gmail.com

© 2012 Chinese Physical Society and IOP Publishing Ltd

the security of the Noh09 protocol relies on a so-called equivalent EDP (entanglement distillation protocol), which is designed to simulate the function of a counterfactual quantum cryptography (CQC) protocol. Let us make a simple review of this EDP. Alice first prepared N pairs of entanglement states written by

$$|\Psi\rangle_{\rm A} = \frac{1}{\sqrt{2}}[|H\rangle(\sqrt{T}|0\rangle_a|H\rangle_b + {\rm i}\sqrt{R}|H\rangle_a|0\rangle_b) + |V\rangle(\sqrt{T}|0\rangle_a|V\rangle_b + {\rm i}\sqrt{R}|V\rangle_a|0\rangle_b)],$$

she reserved half of the particles of the entanglement states in her own and transmitted the other half to Bob, after the photons pass through the BS. Bob's blocking operation in the Noh09 protocol is abstracted by a unitary operation which is performed on the pulse in path b. Another unitary is performed on the entanglement state to check the consistency, the sifted key is generated on the subset of the entanglement states for which the consistency holds. Then its security under the collective attacks was proved in the rest of the paper, by bounding the phase error rate of the quantum states.

We find that there are some flaws in Yin's proposal with two aspects. First, Eve may not try to entangle her ancilla with the intercepted states, since she has no access into mode a, thus she is still unable to gain any information about the key even if she succeeds in doing this. Note that it is not necessary to entangle her ancilla with the intercepted one for the sake of getting more information. Second, the so-called equivalence between their EDP and the original protocol needs further consideration, ambiguous evidence shows that the outputs from the two protocols differ sometimes. In other words, it is expected to give an equivalence proof from the views of Alice, Bob, and Eve.

In this paper, we give a more intuitive security proof of CQC against the general intercept-resend attacks. The paper is organized as follows. In Section 2, the general intercept-resend attack is introduced, then an attack model is given. Based on this model, the key rate for the ideal protocol is calculated in Section 3. In Section 4, we investigate the practical security of a real-world CQC implementation in a scenario where a time-shift attack based on imperfect detectors is performed. At last, a conclusion is drawn.

2. Modeling the attacks

In a conventional QKD scheme, Eve usually performs a unitary operation on the intecepted qubits and her ancilla, she will not measure her probes until all the classical information is revealed . Thus Eve's attack can be formalized by a unitary operator and an optimal measurement operator conditioned on the classical information. However, this strategy may not account for a CQC protocol, because information carriers are never transmitted via the channel, Eve will always fail to entangle her probe with the right qubits. A more intuitive way is to launch a general intercept-resend attack, with which Eve may skillfully misleed both Alice and Bob as much possible as she can to agree with a key on those particles travelling in path b, and corrupt as much as possible the instances in event $\rm E_1$.

The quantum system is an abstract Hilbert space which includes both actual and fictive subsystems. We denote the initial quantum state in this abstract space by $H_{\rm A}\otimes H_{\rm B}\otimes H_{\rm E}\otimes H_{\rm D}$, where $H_{\rm A}$, $H_{\rm B}$, and $H_{\rm E}$ are the subspaces owned by Alice, Bob, and Eve respectively, and $H_{\rm D}$ is a fictive system which describes the measurement results of detectors $\rm D_1$, $\rm D_2$, and $\rm D_3$. The initial state can then be expressed by

$$|\Phi_0\rangle = \{|\Xi_{xyz}\rangle({\rm i}\sqrt{R}|p\rangle_a|0\rangle_b + \sqrt{T}|0\rangle_a|p\rangle_b)|q\rangle_{\rm B}|e\rangle_{\rm E}\}^{\otimes n}. \quad (1)$$

Here $|\Xi_{xyz}\rangle$ is a fictive state where x, y, and z denote the measurement results of detectors $\rm D_1$, $\rm D_2$, and $\rm D_3$, respectively, p and q are variables set in H or V, and the variable e denoting Eve's result of measurement takes the value from $\{0, H, V\}$. Let $n = 1$, for simplicity, so that only a one-bit key is analyzed, hence equation (1) is changed to

$$|\Phi_0\rangle = |\Xi_{xyz}\rangle({\rm i}\sqrt{R}|p\rangle_a|0\rangle_b + \sqrt{T}|0\rangle_a|p\rangle_b)|q\rangle_{\rm B}|e\rangle_{\rm E}. \quad (2)$$

Before Eve's attack is modeled, it is necessary to define a unitary operator describing the behaviors of the original protocol,

$$\begin{aligned}
&|\Phi_1\rangle \\
&= U_{\rm CQC}|\Xi_{xyz}\rangle({\rm i}\sqrt{R}|p\rangle_a|0\rangle_b + \sqrt{T}|0\rangle_a|p\rangle_b)|q\rangle_{\rm B}|e\rangle_{\rm E} \\
&= \frac{1}{2}\Bigg\{\sum_{\{p,q|q=p\}} {\rm i}\sqrt{R}|p\rangle_a|0\rangle_b|q\rangle_{\rm B}(\sqrt{T}|\Xi_{p00}\rangle + \sqrt{R}|\Xi_{0p0}\rangle) \\
&\quad + \sum_{\{p,q|q=p\}} \sqrt{T}|0\rangle_a|p\rangle_b|q\rangle_{\rm B}|\Xi_{00p}\rangle \\
&\quad + \sum_{\{p,q|q\neq p\}} ({\rm i}\sqrt{R}|p\rangle_a|0\rangle_b \\
&\quad + \sqrt{T}|0\rangle_a|p\rangle_b)|q\rangle_{\rm B}|\Xi_{0p0}\rangle\Bigg\}\otimes|e\rangle_{\rm E}. \quad (3)
\end{aligned}$$

Eve's attack is now written by another unitary operator $U_{\rm E}$ acting on the Hilbert space $H_{\rm A}\otimes H_{\rm B}\otimes H_{\rm E}\otimes H_{\rm D}$,

$$
\begin{aligned}
|\Phi_{\rm out}\rangle &= U_{\rm E}|\Phi_1\rangle \\
&= \frac{1}{2}\sum_{\{p,q|q=p\}} {\rm i}\sqrt{RT}|p\rangle_a|0\rangle_b|q\rangle_b\Big(\sum_{\{x,y,z|x=p,y\cup z\neq 0\}}\alpha^1_{xyz}|\Xi_{xyz}\rangle|0\rangle_{\rm E}+\alpha^1_{p00}|\Xi_{p00}\rangle|0\rangle_{\rm E}\Big) \\
&\quad+\frac{1}{2}\sum_{\{p,q|q=p\}}(-R)|p\rangle_a|0\rangle_b|q\rangle_{\rm B}\Big(\sum_{\{x,y,z|y=p,x\cup z\neq 0\}}\alpha^2_{xyz}|\Xi_{xyz}\rangle|0\rangle_{\rm E}+\alpha^2_{0p0}|\Xi_{0p0}\rangle|0\rangle_{\rm E}\Big) \\
&\quad+\frac{1}{2}\sum_{\{p,q|q\neq p\}}\Big({\rm i}\sqrt{R}|p\rangle_a|0\rangle_b\Big(\sum_{\{x,y,z|x\cup y\neq 0,z\neq 0\}}\alpha^3_{xyz}|\Xi_{xyz}\rangle|0\rangle_{\rm E}+\alpha^3_{p00}|\Xi_{p00}\rangle|0\rangle_{\rm E}+\alpha^3_{0p0}|\Xi_{0p0}\rangle|0\rangle_{\rm E}\Big) \\
&\quad+\frac{1}{2}\sum_{\{p,q|q=p\}}\sqrt{T}|0\rangle_a|p\rangle_b\Big)|q\rangle_{\rm B}(\alpha^4_{p00}|\Xi_{p00}\rangle|p\rangle_{\rm E}+\alpha^4_{0p0}|\Xi_{0p0}\rangle|p\rangle_{\rm E}+\alpha^4_{00p}|\Xi_{00p}\rangle|p\rangle_{\rm E}) \\
&\quad+\frac{1}{2}\sum_{\{p,q|q\neq p\}}\sqrt{T}|0\rangle_a|p\rangle_b)|q\rangle_{\rm B}(\alpha^5_{p00}|\Xi_{p00}\rangle|p\rangle_{\rm E}+\alpha^5_{0p0}|\Xi_{0p0}\rangle|p\rangle_{\rm E}+\alpha^5_{00q}|\Xi_{00q}\rangle|p\rangle_{\rm E}). \qquad (4)
\end{aligned}
$$

This model is naturally obtained according to the general intercept-resend strategy, and five cases need to be focused on: (C1) Alice's polarization is consistent with Bob's, mode b is in a vacuum and detector $\rm D_1$ is supposed to click. In this case, Eve's attack causes at least two detectors to click at the same time, this explains the term

$$\sum_{\{x,y,z|x=p,y\cup z\neq 0\}}\alpha^1_{xyz}|\Xi_{xyz}\rangle|0\rangle_{\rm E},$$

the other term $\alpha^1_{p00}|\Xi_{p00}\rangle|0\rangle_{\rm E}$ represents the probability of those uncorrupted events. In particular, $|0\rangle_{\rm E}$ denotes the state when Eve gets nothing about the key. (C2) Alice's polarization is consistent with Bob's, mode b is in a vacuum and detector $\rm D_2$ is supposed to click. Eve's attack is then presented in the term

$$\sum_{\{x,y,z|y=p,x\cup z\neq 0\}}\alpha^2_{xyz}|\Xi_{xyz}\rangle|0\rangle_{\rm E}.$$

(C3) Alice's polarization is inconsistent with Bob's, mode b is in a vacuum and detector $\rm D_2$ is supposed to click. Similarly,

$$\sum_{\{x,y,z|x\cup y\neq 0,z\neq 0\}}\alpha^3_{xyz}|\Xi_{xyz}\rangle|0\rangle_{\rm E}$$

describes the events in which at least two detectors click simultaneously, $\alpha^3_{p00}|\Xi_{p00}\rangle|0\rangle_{\rm E}$ or $\alpha^3_{0p0}|\Xi_{0p0}\rangle|0\rangle_{\rm E}$ tells that detector $\rm D_1$ or $\rm D_2$ clicks. (C4) Alice's polarization is consistent with Bob's, mode b is a non-vacuum and detector $\rm D_3$ is supposed to click. In this case, it is possible for Eve to compromise the key, as the term $\alpha^4_{p00}|\Xi_{p00}\rangle|p\rangle_{\rm E}$ says, the explanation to $\alpha^4_{00p}|\Xi_{00p}\rangle|p\rangle_{\rm E}$ is that Eve needs to probe Bob's polarization sometimes to adjust her following interceptions. (C5) Alice's polarization is inconsistent with Bob's, mode b is a non-vacuum and detector $\rm D_2$ is supposed to click. It is similar with the case C4 except that Eve is unable to get any information from her results in this case.

We must point out that this formalization is general, or in other words is valid for any intercept-resend attack, because the coefficients $\{\alpha^i_{xyz}|i=1,2,3,4,5\}$ are arbitrary. In the next section, we will bound Eve's information based on this model, and give a more rigorous key rate of CQC under the condition that all the quantum apparatuses are perfect.

3. Theoretical security against general intercept-resend attacks

In a security proof of conventional quantum key distribution schemes, such as the BB84 protocol, security definition is important and an asymptotic security is ultimately concluded as a function of the length of the quantum sequence. In this paper the infinite-length rather than an asymptotic security is investigated.

Suppose all the quantum apparatuses are perfect, especially the detectors, of which the efficiency η reaches 1 and the dark counter rate $p_{\rm d}$ is set to zero, any event in which more than one detector clicks is definitely absent, because it implies the existence of an eavesdropper Eve. Thereby the model is simplified after fixing the terms as $\{\alpha^1_{xyz}=0|x=p,y\cup z\neq 0\}$, $\{\alpha^2_{xyz}=0|y=p,x\cup z\neq 0\}$, and $\{\alpha^3_{xyz}=0|z\cup y\neq 0,z\neq 0\}$. Then one obtains

$$
\begin{aligned}
&|\Phi_{\rm out}\rangle \\
&= U_{\rm E}|\Phi_1\rangle
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{\{p,q|q=p\}} \mathrm{i}\sqrt{RT}|p\rangle_a|0\rangle_b|q\rangle_{\mathrm{B}}\alpha^1_{p00}|\Xi_{p00}\rangle|0\rangle_{\mathrm{E}} \\
&+ \frac{1}{2}\sum_{\{p,q|q=p\}} (-R)|p\rangle_a|0\rangle_b|q\rangle_{\mathrm{B}}\alpha^2_{0p0}|\Xi_{0p0}\rangle|0\rangle_{\mathrm{E}} \\
&+ \frac{1}{2}\sum_{\{p,q|q\neq p\}} \Big(\mathrm{i}\sqrt{R}|p\rangle_a|0\rangle_b(\alpha^3_{p00}|\Xi_{p00}\rangle|0\rangle_{\mathrm{E}} \\
&+ \alpha^3_{0p0}|\Xi_{0p0}\rangle|0\rangle_{\mathrm{E}}) \\
&+ \frac{1}{2}\sum_{\{p,q|q=p\}} \sqrt{T}|0\rangle_a|p\rangle_b\Big)|q\rangle_{\mathrm{B}}(\alpha^4_{p00}|\Xi_{p00}\rangle|p\rangle_{\mathrm{E}} \\
&+ \alpha^4_{0p0}|\Xi_{0p0}\rangle|p\rangle_{\mathrm{E}} + \alpha^4_{00p}|\Xi_{00p}\rangle|p\rangle_{\mathrm{E}}) \\
&+ \frac{1}{2}\sum_{\{p,q|q\neq p\}} \sqrt{T}|0\rangle_a|p\rangle_b)|q\rangle_{\mathrm{B}}(\alpha^5_{p00}|\Xi_{p00}\rangle|p\rangle_{\mathrm{E}} \\
&+ \alpha^5_{0p0}|\Xi_{0p0}\rangle|p\rangle_{\mathrm{E}} + \alpha^5_{00q}|\Xi_{00q}\rangle|p\rangle_{\mathrm{E}}). \qquad (5)
\end{aligned}
$$

This model is correct, for instance in the ideal case where there is no eavesdropper one may obtain

$$p_{\mathrm{D}_1} = \frac{RT}{2}, \qquad (6)$$

$$p_{\mathrm{D}_2} = \frac{R^2}{2} + \frac{1}{2}, \qquad (7)$$

$$p_{\mathrm{D}_3} = \frac{T}{2}, \qquad (8)$$

$$p^1_e = p^2_e = p^3_e = 0. \qquad (9)$$

Note that any pair of inconsistent measurement results of Alice and Bob indicates an error bit, it is easy to confirm six important parameters which are crucial to bound Eve's information I_{AE},

$$
\begin{aligned}
p_{\mathrm{D}_1} &= \frac{RT}{2}|a^1_{p00}|^2 + \frac{R}{2}|a^3_{p00}|^2 \\
&+ \frac{T}{2}|a^4_{p00}|^2 + \frac{T}{2}|a^5_{p00}|^2, \qquad (10)
\end{aligned}
$$

$$
\begin{aligned}
p_{\mathrm{D}_2} &= \frac{R^2}{2}|a^2_{0p0}|^2 + \frac{R}{2}|a^3_{0p0}|^2 \\
&+ \frac{T}{2}|a^4_{0p0}|^2 + \frac{T}{2}|a^5_{0p0}|^2, \qquad (11)
\end{aligned}
$$

$$p_{\mathrm{D}_3} = \frac{T}{2}(|a^4_{00p}|^2 + |a^5_{00q}|^2), \qquad (12)$$

$$p^1_e = \frac{R}{2}|a^3_{p00}|^2 + \frac{T}{2}|a^5_{p00}|^2, \qquad (13)$$

$$p^2_e = 0, \qquad (14)$$

$$p^3_e = \frac{T}{2}|a^5_{00q}|^2. \qquad (15)$$

Now the key rate, as well as the bound of Eve's information, can be given by the following theorem.

Theorem 1 For a given observable probability distribution $\{P_{\mathrm{D}_1}, P_{\mathrm{D}_2}, P_{\mathrm{D}_3}\}$ and the error rates P^1_{e}, P^2_{e}, and P^3_{e}, the mutual information between Alice and Bob is

$$I_{\mathrm{AB}} = p^*_{\mathrm{D}_1}\left[1 - h\left(\frac{p^1_{\mathrm{e}}}{p_{\mathrm{D}_1}}\right)\right],$$

and between Alice and Eve is

$$I_{\mathrm{AE}} = \left(\frac{T}{2} + p^3_{\mathrm{e}} - p_{\mathrm{D}_3}\right)R,$$

hence the infinite-effect key rate is

$$
\begin{aligned}
m_{\mathrm{k}} &= I_{\mathrm{AB}} - I_{\mathrm{AE}} = p_{\mathrm{D}_1} - \left(\frac{T}{2} + p^3_{\mathrm{e}} - p_{\mathrm{D}_3}\right)R \\
&- p_{\mathrm{D}_1} * h\left(\frac{p^1_{\mathrm{e}}}{p_{\mathrm{D}_1}}\right). \qquad (16)
\end{aligned}
$$

Proof See Appendix A.

So far we have proved the security of CQC against the general intercept-resend attacks in the context that quantum apparatus works according to its specification. As we can see in Eq. (16), the key rate is inherently bounded by the probabilities of events E_1, E_2, and E_3, and the error rates in events E_1 and E_3 (Eve's attack does not affect event E_2, as seen from Eq. (14)). This coincides with the real situation in the sense that any abnormal probability distribution or a relatively high error rate in one of the events immediately implies the existence of an eavesdropper. Also note that any protocol is secure only if it satisfies $m_k > 0$. Fortunately, as shown in Fig. 1, a positive security key rate can be achieved when the error rates in events E_1 and E_3 are small enough and the probability distribution is not far from normal, i.e., $p_{\mathrm{D}_1} = RT/2$, $p_{\mathrm{D}_2} = R^2/2 + 1/2$, and $P_{\mathrm{D}_3} = T/2$.

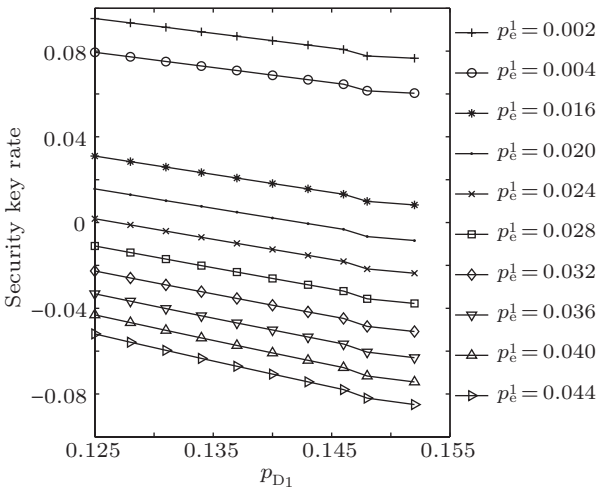


Fig. 1. The theoretical key rate of CQC as a function of P_{D_1}. The curves are obtained from Eq. (16), given randomly chosen vectors $\boldsymbol{p}_{\mathrm{D}_1}$ = (0.125, 0.128, 0.131, 0.134, 0.137, 0.14, 0.143, 0.146, 0.148, 0.152) and $\boldsymbol{p}_{\mathrm{D}_3}$ = (0.25, 0.24, 0.23, 0.22, 0.21, 0.20, 0.19, 0.18, 0.17, 0.16). Note that the values in these vectors are carefully selected to keep them self-contained. For instance, $\boldsymbol{p}_{\mathrm{D}_1}$ is monotonously ascending, consequently, $\boldsymbol{p}_{\mathrm{D}_3}$ descends monotonously. They are indeed in agreement with the data in a practical case.

4. A time-shift attack in practical CQC system

We have been reminded that imperfect detectors can be the loophole in practical QKD systems,[16] several elaborately conceived attacks can be applied to crack the whole system without detection.[17−19] Unfortunately, such a problem also exists in practical CQC implementations. In this section, it will be shown that practical CQC systems are vulnerable to device imperfections by an example of a time-shift attack based on imperfect detector efficiency.

Due to the dark counter rate, two or three detectors might click simultaneously, this event is denoted by E_4. Assume that Eve is only as powerful as allowed by the constraints of quantum physics. Naturally, she must lower the probability P_{E_4} to make herself undetectable, thus it should satisfy

$$P_{\mathrm{E}_4} \leq 2P_{\mathrm{d}}, \tag{17}$$

i.e.,

$$\begin{aligned} p_{\mathrm{D}4} = {} & \frac{RT}{2} \sum_{\{x,y,z|x=p,y\cup z\neq 0\}} |\alpha^1_{xyz}|^2 \\ & + \frac{R^2}{2} \sum_{\{x,y,z|y=p,x\cup z\neq 0\}} |a^2_{xyz}|^2 \\ & + \frac{R}{2} \sum_{\{x,y,z|x\cup y\neq 0,z\neq 0\}} |a^3_{xyz}|^2 \leq 2P_d. \end{aligned} \tag{18}$$

Event E_4 is advantageous to Eve, yet as it is bounded by the dark counter rate, the corrupted information seems trivial, otherwise Eve takes high risk of being detected. Lemma 1 gives the maximal corrupted information by Eve.

Lemma 1 For a given dark counter rate p_{d} of the detector, the maximal corrupted bit rate is obtained by

$$\begin{aligned} r_{\mathrm{c}\max} = {} & (T+1)P_{\mathrm{d}} + [p_{\mathrm{D}_1} - (T+1)P_{\mathrm{d}}]h \\ & \times \left(\frac{p^1_{\mathrm{e}}}{p_{\mathrm{D}_1} - (T+1)P_{\mathrm{d}}}\right) \\ & - p_{\mathrm{D}_1} h\left(\frac{p^1_{\mathrm{e}}}{p_{\mathrm{D}_1}}\right). \end{aligned} \tag{19}$$

Proof See Appendix B.

Another important parameter of the detector is the efficiency η, which brings the risk: Eve may artificially manipulate her device to lower the efficiency η, in order to pass the test with a greater probability. To compromise the key, Eve has to probe Bob's polarization by sending fake photons, then these photons are unlikely to be caught by detector D_3 if the efficiency of detector D_3 is low enough. Therefore Eve is able to control an imperfect detector simply by using an optical delay, i.e., she may launch a time-shift attack[18] to artificially lower the efficiency to an acceptable level, due to the distinguishability between the real efficiency function (refer to Fig. 2) and the ideal one. This evidently causes an increment on the mutual information between Alice and Eve, demonstrated by Lemma 2.

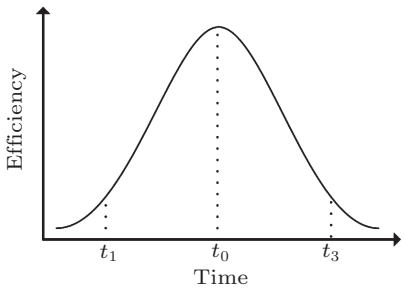


Fig. 2. Efficiency of an imperfect detector as a function of time.

Lemma 2 Given a prior fixed detector efficiency η, the increment of the mutual information between Alice and Eve is

$$\triangle I^{\eta}_{\mathrm{AE}} = \frac{1-\eta}{\eta}(p_{\mathrm{D}_3} - p^3_{\mathrm{e}})R. \tag{20}$$

Proof See Appendix C.

Combined with Lemmas 1 and 2, the total of the reduced information m_{k} is immediately obtained by a sum of the two,

$$\triangle m_{\mathrm{k}} = \gamma_{\mathrm{c}\max} + \triangle I^{\eta}_{\mathrm{AE}}. \tag{21}$$

So the key rate under a time-shift attack is changed to

$$m'_{\mathrm{k}} = m_{\mathrm{k}} - \triangle m_{\mathrm{k}}. \tag{22}$$

As shown in Fig. 3, a negative security key rate can be achieved. Consequently, Eve can obtain full information about the key even if $m_{\mathrm{k}} > 0$. A three-dimensional illustration of the key rate as functions of P_{D_1} and P_{D_3} is shown in Fig. 4 for a better understanding of this conclusion. Since there may be, for most of the QKD protocols, unexpected difficulties when we try to fill the gap between the ideal model and a realistic implementation to a negligible level at present, it is expected that more applicable CQC protocols are required in the future.

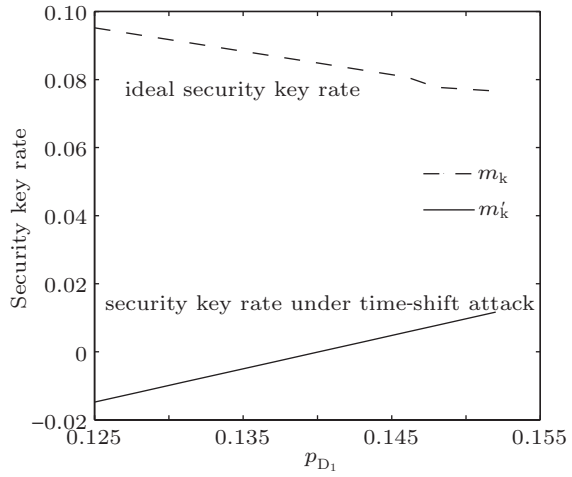


Fig. 3. The key rate under a time-shift attack.

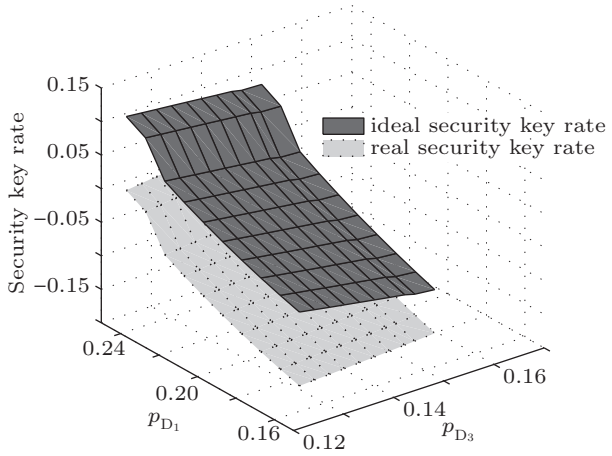


Fig. 4. The security key rate under time-shift attack as functions of P_{D_1} and P_{D_3}, compared to the ideal one.

5. Conclusion

We have proved the security of counterfactual quantum cryptography against general intercept-resend attacks, by modeling it with two unitary operations, which are abstracted from the Noh09 protocol and the attack strategy respectively. Note that collective attacks in conventional QKD protocols do not account for CQC equally, yet intercept-resend attack seems to be more advantageous to the eavesdropper. Intuitively, Eve may corrupt the key-to-be events and mislead the key generation. The protocol is proved to be secure against such attacks in the context that quantum apparatuses are perfect, then the security key rate is bounded by the probability distribution and the error rates of three intrinsic events. Our result is complementary to the one in Ref. [14], in which the error rate in events D_3 is not involved in the conclusion. Also note that Eve's attack not only alters the probability distribution, but also introduces errors, thus, our proof coincides with the real situation more smoothly. Being demonstrated by the example of a time-shift attack, it is also shown that practical CQC implementation is vulnerable to imperfect quantum apparatuses, this opens another problem how to put forward a secure CQC system in real life.

Acknowledgement

The authors are grateful to professor Lev Vaidman for informing us about the importance of interaction-free measurement on which CQC is based.

Appendix A: Proof of Theorem 1

To prove Theorem 1, we start with Eq. (5) and fix an important parameter $\alpha^4_{\rho 00}$, which is given by the unitarity of U_E shown below

$$|\alpha^4_{p00}|^2 + |\alpha^4_{0p0}|^2 + |\alpha^4_{00p}| = 1. \tag{A1}$$

According to case (C4), which is the only one Eve can be successful in, when the fake photon transmitted by Eve is received by Alice, the probabilities for the two probabilities are immediately set by the BS,

$$|\alpha^4_{p00}|^2/|\alpha^4_{0p0}|^2 = R/T, \tag{A2}$$

combining with Eq. (12), (15), (16), and (22), we obtain

$$|\alpha^4_{p00}|^2 = \left(\frac{T}{2} + P^3_e - P_{D_3}\right)\frac{2R}{T}. \tag{A3}$$

With the binary shannon entropy function, Eve's information I_{AE} is therefore bounded by

$$\begin{aligned} I_{AE} &= I(p:e|E_1 = 1, C4 = 1) \\ &= p(E_1 = 1, C4 = 1)[1 - h(p(\bar{e}|p)] \\ &= p(E_1 = 1, C4 = 1) \\ &= \frac{T}{2}|\alpha|^4_{p00} \\ &= \left(\frac{T}{2} + P^3_e - P_{D_3}\right)R, \end{aligned} \tag{A4}$$

where $E_1 = 1$ denote that detector D_1 clicks, $C4 = 1$ means that Eve confirms that case C4 is occupied, and $h(x) = -x\log x - (1-x)\log(1-x)$ is the binary Shannon entropy function. Obviously, Eve obtains no information when p_{D_3} is normal and there is no error in event E_3. Thereby it is natural to obtain that $I_{AE} = 0$ if $P_{D_3} = T/2$ and $P^3_e = 0$.

To bound the mutual information between Alice and Bob, i.e., $I_{\rm AB}$, let us first find the QBER defined by QBER= p(error)/p(arrive), which is different to the one in conventional QKD schemes. Since only event $\rm E_1$ is useful for the key generation, the QBER can be equalized to the error rate for event $\rm E_1$, thus, the actual QBER for CQC protocol is obtained by

$$\mathrm{QBER}_{D_1=1} = p(\mathrm{error}|D_1=1) = \frac{P_{\rm e}^1}{P_{\rm D_1}}. \tag{A5}$$

With the similar technique of bounding Eve's information mentioned above, we can obtain $I_{\rm AB}$

$$I_{\rm AB} I(q:p|D_1=1) = p_{\rm D_1}[1-h(\mathrm{QBER}_{D_1=1})], \tag{A6}$$

where $\mathrm{QBER}_{\rm D_1=1}$ is given by Eq. (A5).

In particular, we could, from Eq. (A6), obtain $I_{\rm AB} = RT/2$ given $\mathrm{QBER}_{D_1=1} = 0$. This is consistent with the result for the ideal case. Now the key rate is naturally obtained by

$$\begin{aligned} m_{\rm k} &= I_{\rm AB} - I_{\rm AE} \\ &= P_{\rm D_1} - \left(\frac{T}{2} + P_{\rm e}^3 - P_{\rm D_3}\right) R \\ &\quad - P_{\rm D_1} h\left(\frac{P_{\rm e}^1}{P_{\rm D_1}}\right). \end{aligned} \tag{A7}$$

This concludes the proof.

Appendix B: Proof of Lemma 1

Before we start the proof, a discussion about Eq. (4) is presented. It is shown that event $\rm E_4$ occurs only when mode b is in vacuum, and Eve cannot distinguish cases C1, C2, and C3, her corruption to the key-to-be events in case C1 absolutely affects the other two. It is not difficult to find that the effect to the three cases is supposed to be the same as a result of Eve's corruption, then we obtain

$$\begin{aligned} &\sum_{\{x,y,z|x=p,y\bigcup z\neq 0\}} |\alpha_{xyz}^1|^2 \\ &= \sum_{\{x,y,z|y=p,x\bigcup z\neq 0\}} |\alpha_{xyz}^2|^2 \\ &= \sum_{\{x,y,z|x\bigcup y\neq 0,z\neq 0\}} |\alpha_{xyz}^3|^2. \end{aligned} \tag{B1}$$

To reach the maximum, the total probability for event $\rm E_4$ is set to be

$$P_{\rm E_4} = 2P_{\rm d}. \tag{B2}$$

From Eqs. (4), (B1), and (B2), the following expression can be confirmed

$$\begin{aligned} &\sum_{\{x,y,z|x=p,y\bigcup z\neq 0\}} |\alpha_{xyz}^1|^2 \\ &= \sum_{\{x,y,z|y=p,x\bigcup z\neq 0\}} |\alpha_{xyz}^2|^2 \\ &= \sum_{\{x,y,z|x\bigcup y\neq 0,z\neq 0\}} |\alpha_{xyz}^3|^2 = \frac{2}{R} P_{\rm d}. \end{aligned} \tag{B3}$$

In addition, from Theorem 1 the mutual information between Alice and Bob when the corruption occurs is

$$I_{\rm AB}^{\rm real} = P_{\rm D_1}^{\rm real}\left[1 - h\left(\frac{P_{\rm ereal}^1}{P_{\rm D_1}^{\rm real}}\right)\right], \tag{B4}$$

where $P_{\rm D_1}^{\rm real}$ denotes the real probability for event $\rm E_1$ and $P_{\rm ereal}^1$ is the corresponding error rate. Next, we consider the ideal case, correspondingly, the two parameters are changed to

$$\begin{aligned} P_{\rm D_1}^{\rm real} = P_{\rm D_1}^{\rm ideal} - \Big(&\frac{RT}{2} \sum_{\{x,y,z|x=p,y\cup z\neq 0\}} |\alpha_{xyz}^1|^2 \\ &+ \frac{R}{2} \sum_{\{x,y,z|x\cup y\neq 0,z\neq 0\}} |\alpha_{xyz}^3|^2\Big), \end{aligned} \tag{B5}$$

$$P_{\rm eideal}^1 = P_{\rm ereal}^1. \tag{B6}$$

Explanation to Eq. (B5) follows: It is assumed the term $\sum_{\{x,y,z|x\bigcup y\neq 0,z\neq 0\}} |\alpha_{xyz}^3|^2$ only affects $|\alpha_{p00}^3|^2$ and cares nothing about $|\alpha_{0p0}^3|^2$, since it is advantageous to Eve. Similarly, $(R/2)\sum_{\{x,y,z|x\cup y\neq 0,z\neq 0\}} |\alpha_{xyz}^3|^2$ should not affect $|\alpha_{p00}^3|^2$. Now the mutual information for the ideal case is given by

$$I_{\rm AB}^{\rm ideal} = P_{\rm D_1}^{\rm ideal}\left[1 - h\left(\frac{P_{\rm ereal}^1}{P_{\rm D_1}^{\rm ideal}}\right)\right]. \tag{B7}$$

Combined with Eqs. (B3), (B4), (B5), and (B6), the total corruption rate $\gamma_{\rm c\ max}$ defined by $\gamma_{\rm c\ max} = I_{\rm AB}^{\rm ideal} - I_{\rm AB}^{\rm real}$ is bounded, i.e., we have

$$\begin{aligned} \gamma_{\rm c\ max} = &I_{\rm AB}^{\rm ideal} - I_{\rm AB}^{\rm real} = (T+1)P_{\rm d} \\ &+ [P_{\rm D1}^{\rm ideal} - (T+1)P_{\rm d}]h \\ &\times \left(\frac{P_{\rm eideal}^1}{P_{\rm D_1}^{\rm ideal} - (T+1)P_{\rm d}}\right) \\ &- P_{\rm D_1} h\left(\frac{P_{\rm eideal}^1}{P_{\rm D_1}^{\rm ideal}}\right). \end{aligned} \tag{B8}$$

Substituting $P_{\rm D_1}^{\rm ideal}$ and $P_{\rm eideal}^1$ with $P_{\rm D_1}$ and $P_{\rm e}^1$ completes the proof.

Appendix C: Proof of Lemma 2

Proof Note that Eve performs a time-shift attack to control the timing of the detector indirectly. In this case, the efficiency is lowered, thereby the probability for event E_3 is reduced to

$$p_{\mathrm{D3}}^{\mathrm{real}} = \eta p_{\mathrm{D3}}^{\mathrm{ideal}}. \tag{C1}$$

P_{ereal}^{3} and P_{eideal}^{3} are defined to denote the probability for the real case and the ideal case respectively. Similarly, the error rate for event E_3 is

$$p_{\mathrm{ereal}}^{3} = \eta p_{\mathrm{eideal}}^{3}. \tag{C2}$$

From Theorem 1, one immediately obtains

$$I_{\mathrm{AE}}^{\mathrm{ideal}} = \left(\frac{T}{2} + P_{\mathrm{eideal}}^{3} - P_{\mathrm{D_3}}^{\mathrm{ideal}}\right) R. \tag{C3}$$

Similarly, Eve's information in a real case is given by

$$I_{\mathrm{AE}}^{\mathrm{real}} = \left(\frac{T}{2} + P_{\mathrm{ereal}}^{3} - P_{\mathrm{D3}}^{\mathrm{real}}\right) R. \tag{C4}$$

Now combined with Eqs. (C1), (C2), and (C3), the increment of the total information of Eve defined by $\triangle I_{\mathrm{AE}}^{\eta} = I_{\mathrm{AE}}^{\mathrm{real}} - I_{\mathrm{AE}}^{\mathrm{ideal}}$ is

$$\triangle I_{\mathrm{AE}}^{\eta} = \frac{1-\eta}{\eta}(P_{\mathrm{D_3}}^{\mathrm{real}} - P_{\mathrm{ereal}}^{3})R. \tag{C5}$$

Substituting the labels $P_{\mathrm{D_3}}^{\mathrm{real}}$ and P_{ereal}^{3} with $P_{\mathrm{D_3}}$ and P_{e}^{3} completes the proof.

References

[1] Bennett C H and Brassard G 1984 *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, December 9–12, 1984, Bangalore, India, p. 175

[2] Ekert A K 1991 *Phys. Rev. Lett.* **67** 661

[3] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 *Rev. Mod. Phys.* **74** 145

[4] Gottesman D, Lo H K, Lütkenhaus N and Preskill J 2004 *Quantum Infor. Comput.* **4** 325

[5] Stucki D, Gisin N, Guinnard O, Robordy G and Zbinden H 2002 *New J. Phys.* **4** 41

[6] Zhao Y, Qi B, Ma X, Lo H K and Qian L 2006 *Phys. Rev. Lett.* **96** 070502

[7] Zhao Y, Qi B, Ma X, Lo H K and Qian L 2006 *Proceedings of IEEE International Symposium of Information Theory* July 9–14, 2006, Seattle, Washington, USA p. 2094

[8] Elliot C, Pearson D and Troxel G 2003 *Proceeding of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications* August 25–29, 2003, Karlsruhe, Germany p. 227

[9] Zhao Y, Qi B, Ma X, Lo H K and Qian L 2006 *Phys. Rev. Lett.* **96** 070502

[10] Xu F X, Chen W, Wang S, Yin Z Q, Zhang Y, Liu Y, Zhou Z, Zhao Y B, Li H W, Liu D, Han Z F and Guo G C 2008 *Chin. Sci. Bull.* **54** 2991

[11] Noh T G 2009 *Phys. Rev. Lett.* **103** 230501

[12] Elitzur A C and Vaidman L 1993 *Found. Phys.* **23** 987

[13] Vaidman L 2007 *Phys. Rev. Lett.* **98** 160403

[14] Yin Z Q, Li H W, Chen W and Han Z F 2010 *Phys. Rev.* A **82** 042335

[15] Shor P W and Preskill J 2000 *Phys. Rev. Lett.* **85** 441

[16] Makarov V, Anisimov A and Skaar J 2006 *Phys. Rev.* A **74** 441

[17] Makarov V 2008 *Quantum Infor. Comput.* **8** 622

[18] Fung C H F, Qi B, Tamaki K and Lo H K 2007 *Quantum Infor. Comput.* **7** 73

[19] Fung C H F, Qi B, Tamaki K and Lo H K 2007 *Phys. Rev.* A **75** 032314