

**Passive Faraday-mirror attack in a practical two-way quantum-key-distribution system**

Shi-Hai Sun, Mu-Sheng Jiang, and Lin-Mei Liang\*

*Department of Physics, National University of Defense Technology, Changsha 410073, People's Republic of China*

(Received 11 May 2011; published 24 June 2011)

The Faraday mirror (FM) plays a very important role in maintaining the stability of two-way plug-and-play quantum key distribution (QKD) systems. However, the practical FM is imperfect, which will not only introduce an additional quantum bit error rate (QBER) but also leave a loophole for Eve to spy the secret key. In this paper we propose a passive Faraday mirror attack in two-way QKD system based on the imperfection of FM. Our analysis shows that if the FM is imperfect, the dimension of Hilbert space spanned by the four states sent by Alice is three instead of two. Thus Eve can distinguish these states with a set of Positive Operator Valued Measure (POVM) operators belonging to three-dimension space, which will reduce the QBER induced by her attack. Furthermore, a relationship between the degree of the imperfection of FM and the transmittance of the practical QKD system is obtained. The results show that the probability that Eve loads her attack successfully depends on the degree of the imperfection of FM rapidly, but the QBER induced by Eve's attack changes slightly with the degree of the FM imperfection.

DOI: [10.1103/PhysRevA.83.062331](https://doi.org/10.1103/PhysRevA.83.062331)

PACS number(s): 03.67.Hk, 03.67.Dd

**I. INTRODUCTION**

Quantum key distribution (QKD) [1] admits two remote parties, known as Alice and Bob, to share an unconditional secret key, and even the eavesdropper (Eve) has ultimate power admitted by the quantum mechanics. Although the unconditional security has been proved for both the ideal system [2,3] and the practical system [4,5] in past years, some assumptions are set to limit Eve's attack strategy or to ignore some imperfections existing in the practical QKD system. Generally speaking, the practical QKD system is imperfect. Any deviation between the standard security analysis and the practical QKD system will leave a loophole for Eve to obtain more information. In the worst case, Eve can exploit all these imperfections together to maximize her information about the secret key. Thus it is important to do research on the practical QKD system carefully and close these loopholes to guarantee the unconditional security of the key. In fact, some potential attacks using the imperfection of a practical QKD system have been discovered, for example, the timing side channel attack [6], a faked states attack [7], the blinding attack [8], Trojan horse attacks [9], time-shifted attack [10,11], and the phase-remapping attack [12,13]. Therefore when the QKD system is used in practical situations, the legitimate parties should consider the potential attack according to any imperfection existing in the practical system and find defense strategies against them.

In all the practical QKD systems based on long-distance fibers, the major difficulty is to maintain the stability and compensate the birefringence of fiber. In order to resolve this problem, Muller *et al.* proposed an interesting two-way plug-and-play scheme [14], which can compensate the birefringence automatically. In this system Bob sends a strong reference pulse to Alice. Then Alice encodes her information to the reference pulse, attenuates it to the single-photon level, and sends it back to Bob. Since the pulse travels back and forth in the quantum channel, the birefringence is compensated

automatically. However, since Alice admits the pulse into and out of her zone, it will leave a back door for Eve to implement a variable Trojan horse attack [9–13].

In this paper we propose a passive Faraday mirror (PFM) attack in a two-way plug-and-play QKD system based on the imperfection of the Faraday mirror (FM), which plays a very important role in compensating the birefringence of fiber. Our results show that for the Bennett-Brassard 1984 (BB84) protocol [1], when the FM deviates from an ideal situation, the dimension of Hilbert space spanned by the states sent by Alice is three instead of two. Thus it will give Eve more information to spy the secret key. When the legitimate parties are unaware of our attack, the unconditional security of the generated key must be compromised. Thus in practical situations, it is very important for Alice and Bob to consider our attack when they judge whether the plug-and-play QKD system is secure or not.

In the following discussion we first introduce the imperfection of FM and analyze a PFM attack based on this imperfection in Secs. II and III, respectively. In Sec. IV we find the minimal quantum bit error rate (QBER) between Alice and Bob induced by Eve when she uses an optimal and suboptimal Positive Operator Valued Measure (POVM) measurement strategy to implement the intercept-and-resend attack. In Sec. V we give a brief conclusion of this paper.

**II. THE IMPERFECTION OF THE FARADAY MIRROR**

In this section we briefly introduce the two-way plug-and-play system and show why the FM can be used to compensate the birefringence of fiber. Then we show how the imperfection of FM can be used by Eve to spy the secret key.

**A. Plug-and-play QKD system**

A simple diagram of the plug-and-play system [14] without Eve is shown in Fig. 1(a). Bob sends a strong reference pulse to Alice, which is horizontally polarized. The pulse will be divided equally into two parts by a beam splitter (BS), noted as  $a$  and  $b$ . A polarization controller (not shown in Fig. 1) is used to change the polarization of  $b$  to guarantee it can

\*nmliang@nudt.edu.cn

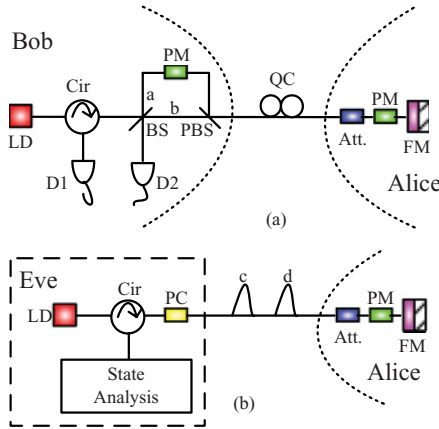


FIG. 1. (Color online) The simple diagram of the plug-and-play system [14] and Eve's attack. LD, laser diode; Cir, circulator; BS, beam splitter; PBS, polarization beam splitter; Att., attenuator; PM, phase modulator; FM, Faraday mirror; QC, quantum channel; D1 and D2, single photon detector; and PC, polarization controller. Part(a) shows plug-and-play system without Eve. Part(b) shows Eve's PFM attack. In the diagram we show only how Eve obtains the information of Alice.  $c$  and  $d$  are two time modes sent to Alice by Eve, and we assume only mode  $c$  is modulated by Alice in this paper.

pass the polarization beam splitter (PBS) totally. Generally speaking, due to the birefringence of fiber, the polarization of  $a$  and  $b$  are random when they arrive at Alice's zone sequentially. However, an FM can be used to compensate the birefringence of channel automatically. When  $a$  and  $b$  return to Bob's zone, their polarizations are orthogonal to that of their initial state. Then they will travel the other path and interfere in BS. Therefore, FM plays an important role in compensating the birefringence of fiber. Now we show why the FM can do this.

The FM is a combination of a  $\theta$  Faraday rotator and an ordinary mirror. In an ideal situation,  $\theta = 45^\circ$  and the Jones matrix of FM can be written as

$$\begin{aligned} \text{FM}(45^\circ) &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\ &= - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \end{aligned} \quad (1)$$

Thus the polarization of the outgoing state is always orthogonal to that of the incoming state, regardless of the input polarization state. It is easy to prove that for any birefringence medium the following equation always holds, that is,

$$T(-\theta') \times \text{FM}(45^\circ) \times T(\theta') = e^{i(\varphi_o + \varphi_e)} \text{FM}(45^\circ), \quad (2)$$

where  $T(\theta')$  and  $T(-\theta')$  are the Jones matrices of birefringence medium when the photon travels forward and backward of the quantum channel, which is given by

$$\begin{aligned} T(\pm\theta') &= \begin{bmatrix} \cos(\theta') & \mp \sin(\theta') \\ \pm \sin(\theta') & \cos(\theta') \end{bmatrix} \begin{bmatrix} \exp(i\varphi_o) & 0 \\ 0 & \exp(i\varphi_e) \end{bmatrix} \\ &\times \begin{bmatrix} \cos(\theta') & \pm \sin(\theta') \\ \mp \sin(\theta') & \cos(\theta') \end{bmatrix}, \end{aligned} \quad (3)$$

where  $\varphi_o$ ,  $\varphi_e$  are the propagation phases of ordinary and extraordinary rays, and  $\theta'$  is the rotation angle between the

reference basis and the eigenmode basis of the birefringence medium. Equation (2) shows clearly that in the ideal situation, the plug-and-play system can compensate the birefringence of medium automatically.

Note that although the plug-and-play system will suffer from the "untrusted source" problem in which the source of the incoming Alice zone is totally controlled by Eve, its security has been rigorously proved in recent works [15–20]. Thus this problem is not considered in this paper and the additional setups for Alice to monitor the "untrusted source" are not shown in Fig. 1.

## B. PFM attack

In the discussion above, we have shown that in the ideal situation, FM can be used to compensate the birefringence of fiber automatically. However, in a practical case, the angle of the Faraday rotator in FM is not exact  $45^\circ$ . Then Eq. (1) is not valid and the Jones matrix of a practical FM should be rewritten as

$$\begin{aligned} \text{FM}(\theta) &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\ &= \begin{bmatrix} \cos(2\theta) & -\sin(2\theta) \\ -\sin(2\theta) & -\cos(2\theta) \end{bmatrix} \\ &= - \begin{bmatrix} \sin(2\varepsilon) & \cos(2\varepsilon) \\ \cos(2\varepsilon) & -\sin(2\varepsilon) \end{bmatrix} \equiv \text{FM}(\varepsilon), \end{aligned} \quad (4)$$

where  $\theta = \pi/4 + \varepsilon$  is the angle of the Faraday rotator in a practical FM. Generally speaking,  $\varepsilon$  is small. For example, in the center wavelengths 1550 and 1310 nm, the maximal rotation angle tolerance is  $1^\circ$  (at  $23^\circ\text{C}$ ) for the popular FM produced by Newport Corp. [21] and General Photonics Corp. [22]. Thus in this paper we consider only the case that  $|\varepsilon| \leq 1^\circ$ .

When FM is imperfect, the birefringence of fiber cannot be compensated totally and additional QBER will be induced. However, the additional QBER is just the minor bug of the practical FM, since  $\varepsilon$  is very small. The major issue is that the imperfection of FM will leave a loophole for Eve to spy the secret key. In the following we show how Eve can use this imperfection, which is called PFM attack in this paper.

The PFM attack is shown in Fig. 1(b). In the diagram we draw only the major part of Eve's attack that shows how Eve probes Alice's information. In order to do this, Eve sends two time modes  $c$  and  $d$  to Alice. Note that the two modes should be coherent, which can be obtained by splitting a pulse with a BS, like the generation of mode  $a$  and  $b$  sent by Bob. The polarization of the two modes should also be the same, which is controlled by Eve's polarization controller (PC). We assume the incoming polarization state  $i$  of the photon that was sent to Alice by Eve is given by

$$i = [\alpha \ \beta]^T. \quad (5)$$

Note that the polarization of the incoming state is totally controlled by Eve; thus  $\alpha$  and  $\beta$  are any complex number that satisfy  $|\alpha|^2 + |\beta|^2 = 1$ . Simply, we consider only the special case that Eve sets  $\alpha = 1$  and  $\beta = 0$  in this paper.

It is easy to prove that the Jones vectors of output polarization state for the two time modes can be written as

$$\begin{aligned} o_c &= \begin{bmatrix} e^{ik\delta} & 0 \\ 0 & 1 \end{bmatrix} \times \text{FM}(\varepsilon) \times \begin{bmatrix} e^{ik\delta} & 0 \\ 0 & 1 \end{bmatrix} \times i \\ &= -e^{ik\delta} \begin{bmatrix} \sin(2\varepsilon)e^{ik\delta} \\ \cos(2\varepsilon) \end{bmatrix} \end{aligned} \quad (6)$$

$$o_d = \text{FM}(\varepsilon) \times i = - \begin{bmatrix} \sin(2\varepsilon) \\ \cos(2\varepsilon) \end{bmatrix} \quad (7)$$

where  $o_c$  and  $o_d$  are the Jones vectors of mode  $c$  and mode  $d$ . Here we assume that only  $c$  is modulated by Alice.  $k = 0, 1, 2, 3$  are the indices of the four states modulated by Alice.  $\delta$  is the phase difference between states, in the standard BB84 scheme,  $\delta = \pi/2$ , but if Eve combines our attack with the phase-remapping attack [12,13],  $\delta \in [0, \pi/2]$ . Note that in Eqs. (6) and (7) we assume that Eve can compensate perfectly the birefringence of Alice's zone by controlling her polarization controller. Although it is still challenging since the birefringence at Alice's side fluctuates, we can ignore this fluctuation since the interval of Eve's photon incoming and outgoing Alice's zone is much smaller than the time for a remarkable fluctuation of the birefringence.

Therefore when FM is imperfect, the states sent by Alice are not the standard BB84 state, which are noted as  $|\phi_k\rangle = (e^{ik\pi/2}|c\rangle + |d\rangle)/\sqrt{2}$ , but four new states that can be written as

$$\begin{aligned} |\Phi_k\rangle &= \frac{1}{\sqrt{2}} \{ \sin(2\varepsilon)e^{i2k\delta}|cH\rangle + \cos(2\varepsilon)e^{ik\delta}|cV\rangle \\ &\quad + \sin(2\varepsilon)|dH\rangle + \cos(2\varepsilon)|dV\rangle \}. \end{aligned} \quad (8)$$

It is easy to check that when  $\varepsilon \neq 0$ , the dimension of Hilbert space spanned by the four new states of Eq. (8) is three. In order to show it clearly, we let  $|H\rangle = \cos(2\varepsilon)|H'\rangle + \sin(2\varepsilon)|V'\rangle$  and  $|V\rangle = -\sin(2\varepsilon)|H'\rangle + \cos(2\varepsilon)|V'\rangle$ , and note  $|cH'\rangle = |e_0\rangle$ ,  $|cV'\rangle = |e_1\rangle$ ,  $|dV'\rangle = |e_2\rangle$ . Note that  $\varepsilon$  can be known by Eve exactly, and thus the transformation from  $\{H, V\}$  to  $\{H', V'\}$  can be implemented by Eve. Then Eq. (8) can be rewritten as

$$\begin{aligned} |\Phi_k\rangle &= \frac{1}{\sqrt{2}} \{ \sin(2\varepsilon)\cos(2\varepsilon)(e^{i2k\delta} - e^{ik\delta})|e_0\rangle \\ &\quad + [\sin^2(2\varepsilon)e^{i2k\delta} + \cos^2(2\varepsilon)e^{ik\delta}]|e_1\rangle + |e_2\rangle \}. \end{aligned} \quad (9)$$

Equations (8) and (9) clearly show that the information encoded by Alice appears not only in the time mode ( $c$  and  $d$ ) but also in the polarization mode ( $H$  and  $V$ ). Obviously, Eve can obtain more information from the four new states than the four standard BB84 states. Thus if Alice and Bob are unaware of the imperfection of FM, it will leave a loophole for Eve to spy the information encoded by Alice. Here we remark that although the true states sent by Alice are shown by Eq. (8), Eve cannot get more information from Eq. (8) than by Eq. (9), since the transformation from Eq. (8) to Eq. (9) is just one unitary operation which cannot give her more information. Thus in the following we use Eq. (9) but not Eq. (8) to continue our discussion.

### III. PFM ATTACK BASED ON AN INTERCEPT-AND-RESEND ATTACK

When FM is imperfect, the states sent by Alice are not the standard BB84 states but the states shown by Eq. (9). Thus Eve can use the operators belonging to three-dimensional Hilbert space to measure the four states. Note that for the case that FM is perfect, the operators that can be used by Eve should be limited in two-dimensional Hilbert space. Generally speaking, when a new imperfection of the system is found by Eve, she can combine all imperfections of the system and attack strategies to maximize her information of the key. However, in this paper we consider only the intercept-and-resend attack, which clearly shows that the generated key will be compromised due to the imperfection of FM.

We consider the following attack: Eve intercepts each pulse returned from Alice's zone and measures it with five POVM operators  $\{M_{\text{vac}}, M_i | i = 0, 1, 2, 3\}$  which satisfy the condition that  $M_{\text{vac}} + \sum_{i=0}^3 M_i = I$ . When Eve obtains the outcome corresponding to  $M_i$ , she resends a standard BB84 state  $|\phi_i\rangle$  to Bob. Here we add an additional POVM operator  $M_{\text{vac}}$ , since Eve does not need to resend all the pulses to Bob due to the loss of channel between Alice and Bob. Thus when she obtains the outcome corresponding to this operator, she blocks the pulse and sends a vacuum state to Bob. Obviously, the above attack is the same as the general intercept-and-resend attack. However, in our attack the dimension of  $M_i$  and  $M_{\text{vac}}$  is three instead of two.

Obviously, Eve cannot distinguish the four states  $|\Phi_k\rangle$  certainly, since they are linearly dependent states. In the following we estimate the QBER between Alice and Bob induced by Eve. According to the measurement theory, the conditional probability that Bob obtains  $|\phi_j\rangle$  given that Alice sends the state  $|\Phi_k\rangle$  is given by

$$\begin{aligned} P(j|k) &= \sum_{i=0}^3 P(B = j | E = i) P(E = i | A = k) \\ &= \sum_{i=0}^3 |\langle \phi_j | \Phi_k \rangle|^2 \text{Tr}(M_i \rho_k), \end{aligned} \quad (10)$$

where  $P(B = j | E = i)$  is the conditional probability that Bob obtains  $|\phi_j\rangle$  given that Eve resends the state  $|\phi_i\rangle$ , and  $P(E = i | A = k)$  is the conditional probability that Eve obtains an outcome corresponding to  $M_i$  given that Alice sends  $|\Phi_k\rangle$ ,  $\rho_k = |\Phi_k\rangle\langle\Phi_k|$ . Note the fact that  $|\langle \phi_j | \Phi_k \rangle|^2 = \delta_{i+2,j} + (\delta_{i+1,j} + \delta_{i+3,j})/2$ , where  $\delta_{i,j}$  is the Kronecker  $\delta$  function. Therefore, the QBER between Alice and Bob induced by Eve's attack is given by

$$\begin{aligned} e^B &= \frac{\sum_{k=0}^3 \sum_{j=0, j \neq k}^3 P(j|k)}{\sum_{k=0}^3 \sum_{j=0}^3 P(j|k)} \\ &= \frac{\sum_{i=0}^3 \text{Tr}(M_i L_i)}{\sum_{i=0}^3 \text{Tr}(M_i \rho)}, \end{aligned} \quad (11)$$

where

$$\begin{aligned} L_i &= \frac{1}{2}\rho_{i+1} + \rho_{i+2} + \frac{1}{2}\rho_{i+3} \\ \rho &= \rho_0 + \rho_1 + \rho_2 + \rho_3. \end{aligned} \quad (12)$$

Generally speaking, the optimal strategy for Eve is to find a set of optimal POVM operators to minimize the QBER,  $e_B$ . Although Eq. (11) is the same as Eq. [3] of Ref. [12] formally, the dimension of POVM operators is three in our attack instead of two in Ref. [12]. Furthermore, in Ref. [12] the authors remark that “the constraint  $\sum_{i=0}^3 M_i \leq I$  is not necessary to minimize  $e^B$ , since any solution to this unconstrained problem can always be scaled down sufficiently to satisfy the constraint.” This conclusion is also true in our case. However, we remark that although the constraint  $\sum_{i=0}^3 M_i \leq I$  will not change the QBER, it will bound the probability that Eve obtains a valid outcome successfully, where the valid outcome means that Eve obtains an outcome corresponding to  $M_i$  but not  $M_{\text{vac}}$ . The probability that Eve obtains an outcome successfully is also an important parameter to describe Eve’s attack, which is given by

$$P_{\text{succ}}^E = \frac{1}{4} \sum_{i=0}^3 \text{Tr}(M_i \rho), \quad (13)$$

where the division by four is that Alice sends one of the four states with equal probability.

Therefore, in order to find the optimal strategy for Eve, she needs to solve the optimization program with two penalty functions, which can be written as:

$$\begin{aligned} & \text{minimize } e^B \quad \text{and} \quad \text{maximize } P_{\text{succ}}^E \\ & \text{subject to } M_i \geq 0, M_{\text{vac}} \geq 0, M_{\text{vac}} + \sum_{i=0}^3 M_i = I, \end{aligned} \quad (14)$$

where  $M_i \geq 0$  and  $M_{\text{vac}} \geq 0$  mean the matrix  $M_i$  and  $M_{\text{vac}}$  are positive. In fact Eve cannot obtain the optimal value of the two penalty functions simultaneously, which is shown in the following. Thus, we consider the QBER  $e^B$  as the major object and  $P_{\text{succ}}^E$  as a minor one.

#### IV. SIMULATION

According to the above analysis, the optimal strategy for Eve is given by the solution of Eq. (14). However, it is hard to find the global optimal solution of Eq. (14) [23]. Thus in this paper our analysis follows the method of Fung *et al.* [12], who use this method to analyze the security of the QKD system for the phase-remapping attack. In this section we first introduce the method of Fung and then estimate the QBER between Alice and Bob with a numerical simulation. Although the method considered here is just suboptimal for Eve, it shows clearly that the secrecy of the generated key will be compromised by our PFM attack. The numerical simulations show that under PFM attack, the QBER between Alice and Bob induced by Eve is less than 25%, which is the QBER under the general intercept-and-resend attack. Furthermore, when Eve combines our attack with the phase-remapping attack, the QBER induced by her attack can be less than 11%, which is the maximal tolerable QBER under the collective attack [24].

##### A. The suboptimal strategy for Eve

According to Ref. [12], the suboptimal strategy can be described as follows:

*Step one:* Instead of distinguishing the four states sent by Alice, here Eve only distinguishes  $\rho_0$  from  $\{\rho_1, \rho_2, \rho_3\}$  and  $\rho_3$  from  $\{\rho_0, \rho_1, \rho_2\}$ . It means that, Eve uses three POVM operators  $\{M_{\text{vac}}, M_0, M_3\}$  to measure the states sent by Alice. When she obtains the outcome corresponding to  $M_0$  or  $M_3$ , she resends a standard BB84 state  $|\phi_0\rangle$  or  $|\phi_3\rangle$  to Bob, otherwise, she resends a vacuum state to Bob. It can also be considered as a special case of the intercept-and-resend attack described in Sec. III that  $M_1 = M_2 = 0$ .

*Step two:* Instead of finding the global optimal POVM operators to minimize  $e^B$ , Eve uses the following POVM operators to measure the states sent by Alice, which are given by:

$$\begin{aligned} M_0 &= x\rho^{-1/2}|C_0\rangle\langle C_0|\rho^{-1/2} \\ M_3 &= x\rho^{-1/2}|C_3\rangle\langle C_3|\rho^{-1/2} \end{aligned} \quad (15)$$

where  $|C_0\rangle$  or  $|C_3\rangle$  is the eigenvector of matrix  $\rho^{-1/2}L_0\rho^{-1/2}$  or  $\rho^{-1/2}L_3\rho^{-1/2}$  corresponding to the minimal un-zero eigenvalue  $\lambda_0$  or  $\lambda_3$ ,  $x$  is the maximal real number that ensure the matrix  $M_{\text{vac}} = I - M_0 - M_3$  is positive. Obviously,  $\{M_{\text{vac}}, M_0, M_3\}$  are positive and sum to the identity, thus they are valid POVM operators.

Here we give two remarks about this suboptimal strategy for Eve.

*Remark one:* Although the global solution of Eq. (14) is not obtained in this paper, it does not matter, since our goal is to show the loophole caused by the imperfection of FM, but not to try to give a strict security analysis. Thus if we can find a set of POVM operators that can show this loophole, it is enough. If Eve who has unlimited computation power wants to optimize her strategy, she can solve the optimization program.

*Remark two:* In the suboptimal strategy described above, Bob only obtains bit 0 in one basis and bit 1 in other basis. Glancingly, it will give Alice and Bob a simple method to defeat our attack, since they only need to estimate the count rate of bit 0 and bit 1 for each basis individually. However, we remark that, the method can only defeat our suboptimal strategy, but can not defeat an optimal attack caused by the imperfection of FM. In fact, the suboptimal strategy is just a toy model to show the security loophole induced by the imperfection of FM. This suboptimal strategy is also used by Fung *et al.* in the phase remapping attack to obtain their main conclusion (Fig. (4) of Ref. [12]). If a true Eve exists, she can solve the optimal programme Eq. (14) and find the optimal strategy for her. Then, the probability that Bob obtains bit 0 and bit 1 is equal in two bases.

##### B. Results

Substituting the suboptimal POVM operators Eq. (15) into Eq. (11) and Eq. (13), it is easy to estimate the QBER for Bob,  $e^B$ , and the probability that Eve obtains outcome successfully,  $P_{\text{succ}}^E$ . Since Eve can combine the phase remapping attack [12] with our attack, we set the phase difference  $\delta \in [0, \pi/2]$  in the following discussion. Here, we remark again that the following results are obtained based on the suboptimal strategy described above. If Eve can find the global solution of Eq. (14), she can improve the following results. Furthermore, note that  $\varepsilon = 0$  is a singular point in PFM attack, since, in this point, the dimension

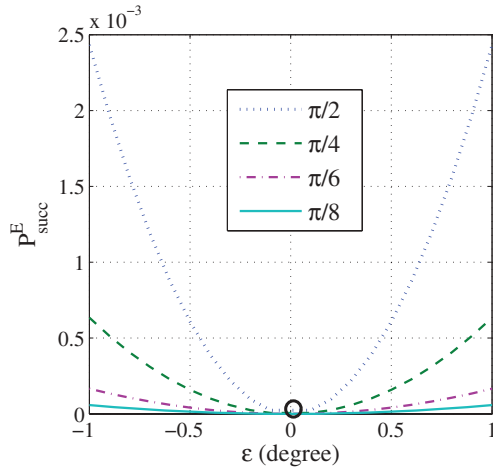


FIG. 2. (Color online) The probability that Eve obtains outcome successfully changes with  $\varepsilon$  for given  $\delta$ . The four lines are obtained by maximizing  $x$  of Eq. (15). Note that the special point that  $\varepsilon = 0$  is unconsidered in our simulation. Since this point is a singular point, in which the dimension of Hilbert space spanned by the four states of Eq. (9) is reduced to two. Thus our attack is invalid in this point.

of Hilbert space spanned by the four states of Eq. (9) is reduced to two. It means that Eve can not implement our attack in this point. Thus, this point is excluded in the following simulation.

Figure 2 shows clearly the probability that Eve obtains outcome successfully,  $P_{\text{succ}}^E$ , changes with  $\varepsilon$  for given  $\delta$ . The larger  $\varepsilon$  is, the easier Eve can load her attack. Here, we remark that  $P_{\text{succ}}^E$  can be explained as the maximal transmittance of channel between Alice and Bob that Eve can load this attack successfully under the suboptimal strategy for given  $\varepsilon$  and  $\delta$  [25]. For example, when  $\varepsilon = 1$  and  $\delta = \pi/2$ ,  $P_{\text{succ}}^E = 2.43 \times 10^{-3}$ . It means that if Eve wants to implement this attack successfully, the transmittance of channel between Alice and Bob should be smaller than  $2.43 \times 10^{-3}$ , which corresponds to a 124 km long fiber (the typical loss of fiber is about 0.21 dB/km). Furthermore,  $P_{\text{succ}}^E$  can also be explained as that, for a given transmittance of channel, Eve can not exploit the imperfection of FM that  $\varepsilon$  is smaller than a given value. For example, when  $\delta = \pi/2$  and  $|\varepsilon| < 0.65^\circ$ ,  $P_{\text{succ}}^E$  is smaller than  $1.029 \times 10^{-3}$ . Thus for a 142.5 km long fiber (the transmittance is about  $1.017 \times 10^{-3}$ ), Eve can not exploit the imperfection of FM that  $|\varepsilon| < 0.65^\circ$ .

The QBER between Alice and Bob induced by Eve's attack is shown in Fig. 3. It shows that even in the case that  $\delta = \pi/2$ , the QBER induced by Eve's attack is much lower than 25% which is QBER induced by the general intercept-and-resend attack. It is also lower than 20%, which is the maximal tolerable QBER in the BB84 protocol under the two-way post-processing [26,27] when Eve does not exploit the imperfection of FM. Furthermore, if Eve combines the phase-remapping attack with our attack, she can reduce the QBER to a very small level. For example, if Eve sets the phase difference  $\delta = \pi/8$ , the QBER induced by her attack is just 3.57%, which is lower than 11% that is the maximal tolerable QBER for the BB84 protocol under the collective attack and one-way post-processing [24]. Then no secret key can be generated when the QBER estimated by Alice and Bob is larger than this

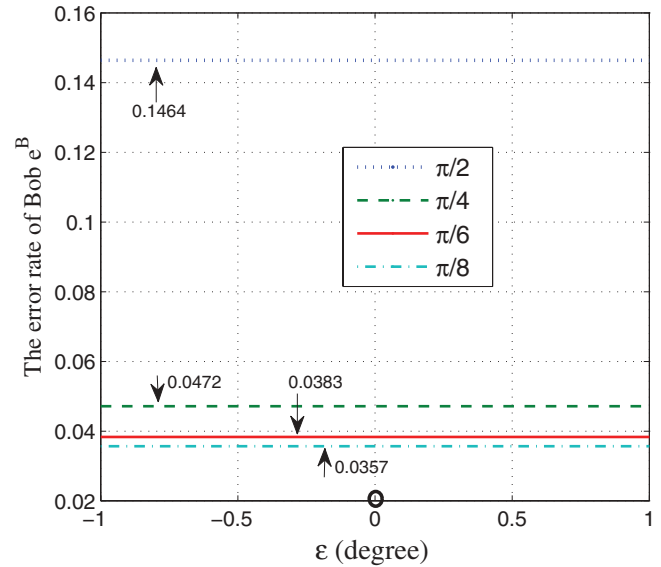


FIG. 3. (Color online) The error rate of Bob changes with  $\varepsilon$  for given  $\delta$ . Note that, due to the same reason as Fig. 2, the special point that  $\varepsilon = 0$  is also unconsidered here. Furthermore, we combine the phase remapping attack with our attack, thus  $\delta$  can be smaller than  $\pi/2$ .

value. Therefore, it is necessary for the legitimate parties to consider our attack in a practical *plug-and-play* QKD system.

It is interesting that, when  $\delta$  is given, the QBER induced by Eve is almost constant and independent of the degree of the imperfection of FM  $\varepsilon$ , which is shown in Fig. 3 clearly. The main reason is that  $\varepsilon$  is very small. In fact,  $e^B$  will change with  $\varepsilon$  slightly, but the difference is so small that it can be ignored. In order to show the conclusion clearly, we consider Eq. (11) with  $o(\varepsilon^2)$ . It is easy to check that, under the suboptimal strategy of Eq. (15),  $e^B = (\lambda_0 + \lambda_3)/2$ , here,  $\lambda_0$  and  $\lambda_3$  are the minimal unzero eigenvalue of  $\rho^{-1/2}L_0\rho^{-1/2}$  and  $\rho^{-1/2}L_3\rho^{-1/2}$ . A simple evolution show that, when  $\delta = \pi/2$ , the three eigenvalues of  $\rho^{-1/2}L_0\rho^{-1/2}$  are  $1/2$ ,  $(1 \pm \sqrt{2}/2)(1 - 2\varepsilon)/2$ . Thus  $\lambda_0 = (1 - \sqrt{2}/2)(1 - 2\varepsilon)/2 = 0.1464 - 0.2929\varepsilon$ . The same result can be obtained for  $\lambda_3$ . Therefore, the error rate of Bob induced by Eve can be written as  $e^B = 0.1464 + o(\varepsilon)$ , which shows clearly that  $e^B$  is constant in order of  $o(\varepsilon)$ . In fact, the strict numerical simulation shows that, for given  $\delta$ ,  $e^B$  is almost constant in order of  $10^{-5}$  for each  $\varepsilon$ . Note that, although  $e^B$  is almost independent of  $\varepsilon$ ,  $\varepsilon$  will affect  $P_{\text{succ}}^E$  obviously (see Fig. 2).

Figure 3 shows that when Eve combines the phase remapping attack with the imperfection of FM, the QBER induced by her attack can be reduced to a very small level. For example, when she sets  $\delta = \pi/4$ , the QBER is just 4.72%. However, when the QBER is reduced, the probability that Eve implements her attack successfully will also be reduced, see Fig. 2. Therefore, when Eve loads her attack, she should make a trade-off between the QBER and the efficiency to maximize her information.

In the following, we compare our attack with the phase remapping attack [12]. The probability that Eve obtains outcome successfully is shown in Fig. 4. It shows that the probability that Eve implements the phase remapping attack

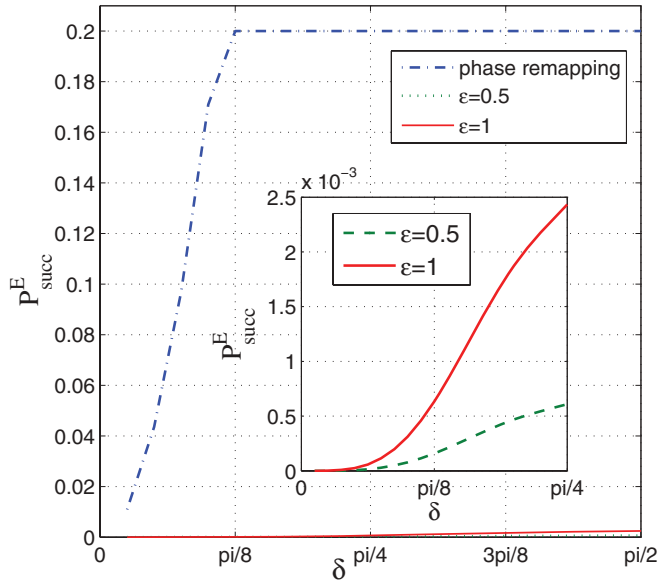


FIG. 4. (Color online) The probability that Eve obtains outcome successfully changes with  $\delta$  for the phase remapping attack and our attack. The red line is obtained for the phase remapping attack according to the method described in Ref. [12]. The blue and green lines are obtained for  $\varepsilon = 1^\circ$  and  $\varepsilon = 0.5^\circ$ , respectively.

successfully is much larger than that of our attack. Although Eve can increase the probability that she implements her attack successfully by increasing the phase difference  $\delta$ , it will induce more QBER which is shown in Fig. 5.

Figure 5 shows that, when the phase difference  $\delta$  is increased, the QBER induced by Eve's attack will increase quickly. This conclusion holds for both the phase remapping attack and our attack. However, the QBER under our attack is much lower than that of the phase remapping attack. For example, when  $\delta = \pi/4$ ,  $e^B = 17.7\%$  for the phase remapping attack, but in our attack,  $e^B = 4.71\%$ .

## V. CONCLUSION

In the two-way *plug-and-play* QKD system, a perfect  $90^\circ$  FM can be used to compensate the birefringence of fiber automatically and perfectly. However, the practical FM is imperfect. Although the deviation from the ideal case is small and the QBER induced by this imperfection is slight, it will leave a loophole for Eve to spy the secret key between Alice and Bob. In fact, when the practical FM deviates from the ideal case, the dimension of Hilbert space spanned by the states sent by Alice is three instead of two. Thus the standard security

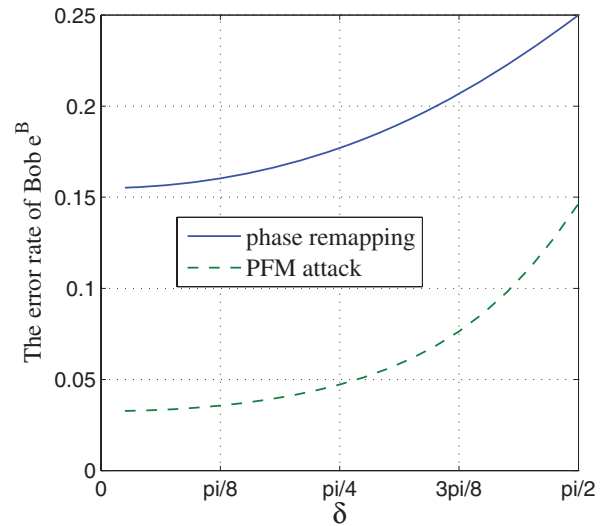


FIG. 5. (Color online) The error rate of Bob changes with  $\delta$  for the phase remapping attack and PFM attack. In the simulation, we set  $\varepsilon = 1^\circ$ .

analysis is invalid here, some careful strategy should be adopted by the legitimate parties to monitor this imperfection.

In this paper, we propose a PFM attack in two way *plug-and-play* QKD system based on the imperfection of FM. The results show that, under this attack, the QBER between Alice and Bob induced by Eve is much lower than 25%, which is the QBER for the general intercept-and-resend attack when FM is perfect. Furthermore, when Eve combines the imperfection of FM with phase remapping attack, the QBER induced by her attack can be lower than 11%, which is the maximal tolerable QBER for the BB84 protocol under the collective attack and one-way post-processing. Therefore, in the practical case, the legitimate parties should pay more attention to the imperfection of FM, otherwise, the secrecy of generated key will be compromised. However, we remark that, although Eve can combine PFM attack with phase-remapping attack to reduce QBER between Alice and Bob induced by her attack, the probability that she can load this attack successfully is dependent on the loss of channel obviously. In other words, Eve can only implement PFM attack in long distance QKD system, which can be estimated in Figs. 2 and 4.

## ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China, Grant No. 61072071. S.-H.S. is supported by Hunan Provincial Innovation Foundation for Postgraduates, No. CX2010B007, and Fund of Innovation, Graduate School of NUDT, No. B100203.

[1] C. H. Bennett and G. Brassard, *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India* (IEEE, New York, 1984), pp. 175–179.

[2] H.-K. Lo and H. F. Chau, *Science* **283**, 2050 (1999).

[3] P. W. Shor and J. Preskill, *Phys. Rev. Lett.* **85**, 441 (2000).

[4] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and John Preskill, *Quantum Inf. Comput.* **4**(5), 325 (2004).

- [5] H. Inamori, N. Lütkenhaus, and D. Mayers, *Eur. Phys. J. D* **41**, 599 (2007).
- [6] A. Lamas-Linares, C. Kurtsiefer, *Opt. Express* **15**, 9388 (2007).
- [7] V. Makarov, Dag R. Hjelm, *J. Mod. Opt.* **52**, 691 (2005).
- [8] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, *Nat. photonics* **4**, 686 (2010).
- [9] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, *Phys. Rev. A* **73**, 022320 (2006).
- [10] Y. Zhao, C. H. F. Fung, B. Qi, C. Chen, and H.-K. Lo, *Phys. Rev. A* **78**, 042333 (2008).
- [11] V. Makarov, A. Anisimov, and J. Skaar, *Phys. Rev. A* **74**, 022313 (2006).
- [12] C. H. F. Fung, B. Qi, K. Tamaki, and H. K. Lo, *Phys. Rev. A* **75**, 032314 (2007).
- [13] F.-H. Xu, B. Qi, and H.-K. Lo, *New J. Phys.* **12**, 113026 (2010).
- [14] A. Muller, T. Herzog, B. Huttner, W. Tittel, H. Zbinden, and N. Gisin, *Appl. Phys. Lett.* **70**, 793 (1997).
- [15] Y. Zhao, B. Qi, H.-K. Lo, and L. Qian, *New J. Phys.* **12**, 023024 (2010).
- [16] Y. Zhao, B. Qi, and H.-K. Lo, *Phys. Rev. A* **77**, 052327 (2008).
- [17] X. Peng, H. Jiang, B.-J. Xu, X.-F. Ma, and H. Guo, *Opt. Lett.* **18**, 2077 (2008).
- [18] B.-J. Xu, X. Peng, and H. Guo, *Phys. Rev. A* **82**, 042301 (2010).
- [19] T. Moroder, M. Curty, and N. Lütkenhaus, *New J. Phys.* **11**, 045008 (2009).
- [20] X.-B. Wang, L. Yang, C.-Z. Peng, and J.-W. Pan, *New J. Phys.* **11**, 075006 (2009).
- [21] [<http://www.newport.com/Fiber-Optic-Faraday-Rotator-Mirrors/835750/1033/catalog.aspx>].
- [22] [<http://www.generalphotonics.com>].
- [23] There are three problems to find the global optimal solution of Eq. (14): (1) There are two penalty functions, it is needed to make a trade off between them, thus it is a problem to deal with it in the numerical simulation. (2) For the five POVM operators belonging to 3-D Hilbert space, the number of variable is 39, it is too much for the simulation. (3) In our problem, the geometrical format of the restrictions is irregular, thus it is hard to find a suitable algorithm to solve this problem. However, it does not matter, since we just want to show the loophole caused by the imperfection of FM, if we can find a set of POVM operators to show this loophole, it is enough.
- [24] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf *et al.*, *Rev. Mod. Phys.* **81**, 1301 (2009).
- [25] Here we associate  $P_{\text{succ}}^E$  with the channel transmittance, but do not consider the transmittance of Bob's setups which includes both the transmittance of Bob's optical setups and the efficiency of Bob's single photon detector (SPD), there are two reasons for this consideration: (1) The channel can be controlled by Eve, but the Bob's setups are placed in Bob's secure zone. Thus Eve can replace the practical channel with a lossless channel, but she cannot change the efficiency of Bob's detectors. (2) In order not to be discovered, Eve should keep Bob's count rate unchanged. When Eve is absent, Bob's count rate is  $C \propto \eta_c \eta_b$ , where  $\eta_c$  is the channel transmittance between Alice and Bob,  $\eta_b$  is the transmittance of Bob's setups. In this paper, we assume Eve resends a single photon pulse to Bob according to her detection result. Thus  $C \propto P_{\text{succ}}^E \eta_b$ , which means  $P_{\text{succ}}^E = \eta_c$ . In fact, we remark that, although Eve cannot change the transmittance of Bob's setups, she can compensate it with a bright light whose intensity is  $\mu$ . Then  $C \propto \mu P_{\text{succ}}^E \eta_b$ . Thus, when  $\mu = 1/\eta_b$ ,  $P_{\text{succ}}^E = \eta_c \eta_b$ , which means Eve also can associate  $P_{\text{succ}}^E$  with the overall efficiency of QKD system by compensating the transmittance of Bob's setups. However, when a bright light is used by Eve, it will introduce other problems, for example, the double click. Therefore, in this paper, we only associate  $P_{\text{succ}}^E$  with the channel transmittance.
- [26] H. F. Chau, *Phys. Rev. A* **66**, 060302(R) (2002).
- [27] K. S. Ranade and G. Alber, *J. Phys. A* **39**, 1701 (2006).