Performance of device-independent quantum key distribution

Zhu Cao, Qi Zhao, and Xiongfeng Ma*

Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China

(Received 8 March 2016; published 13 July 2016)

Quantum key distribution provides information-theoretically-secure communication. In practice, device imperfections may jeopardise the system security. Device-independent quantum key distribution solves this problem by providing secure keys even when the quantum devices are untrusted and uncharacterized. Following a recent security proof of the device-independent quantum key distribution, we improve the key rate by tightening the parameter choice in the security proof. In practice where the system is lossy, we further improve the key rate by taking into account the loss position information. From our numerical simulation, our method can outperform existing results. Meanwhile, we outline clear experimental requirements for implementing device-independent quantum key distribution. The maximal tolerable error rate is 1.6%, the minimal required transmittance is 97.3%, and the minimal required visibility is 96.8%.

DOI: 10.1103/PhysRevA.94.012319

I. INTRODUCTION

Information-theoretical security, which only assumes quantum mechanics, is the core of quantum cryptography. The Bennett–Brassard 1984 (BB84) protocol [1] provides an unconditional secure way of quantum key distribution (QKD), a task that establishes shared keys between two parties [2,3]. The key, combined with one-time pad encryption [4], can then be used to exchange private messages or authenticate messages [5]. The security of QKD protocols often relies on specific models of physical implementation [6]. However, a practical system inevitably deviates from the theoretical model and becomes vulnerable to various side-channel attacks [7–9].

Mayers and Yao [10] first raised the challenge to avoid such attacks by making the devices self-testing. This is also known as the device-independent scenario [11], where the quantum devices are treated as untrusted and uncharacterized. Only a few very reasonable assumptions are made about the devices, such as no direct leakage of local information and reliable random number generation. The self-testing or device-independent quantum key distribution (DIQKD) protocol essentially follows Ekert's QKD protocol [12], which takes advantage of the nonlocality proven in Bell's inequality tests.

Lots of efforts have been devoted to the security proof of DIQKD. For example, Barrett *et al.* [13] analyzed the single round case and Pironio *et al.* [14] analyzed the collective attack case. The challenge was eventually resolved by a seminar work of Vazirani and Vidick [15], where a complete proof of DIQKD without posing any restrictions on Eve's ability was given. Their proof essentially exploits a quantitative version of entanglement monogamy [16]. Later, a different yet also elegant security proof of DIQKD was given by Miller and Shi [17].

In this work, we present clear requirements for experimental devices in a QKD system, including source and measurement devices, such that a secure key can be established by the DIQKD protocol. In addition, we derive an improved analytic key-rate formula based on the work of Vazirani and Vidick. To evaluate its performance, we model an experimental QKD system and compare three postprocessing methods; namely, Vazirani–Vidick, Miller–Shi [17], and ours. Simulation results show that our key rate strictly outperforms the Vazirani–Vidick key rate in all parameter regions and the Miller–Shi key rate in most parameter regions.

The organization of the rest of the paper is as follows: In Sec. II, we review the DIQKD protocol and Vazirani–Vidick proof. In Sec. III, we present our key rate and compare with past results. Section IV puts constraints on the actual parameters of experimental instruments. Finally Sec. V concludes the paper and gives a few outlooks.

II. PRELIMINARIES

A. The Clauser–Horne–Shimony–Holt game

Here we review a basic quantum information concept, the Clauser–Horne–Shimony–Holt (CHSH) game [18]. The setting is illustrated in Fig. 1. Two noncommunicating devices Alice and Bob, enforced by constraints such as spatial separation, take random inputs $x = 0,1$ and $y = 0,1$, respectively and obtain outputs $a = 0,1$ and $b = 0,1$, respectively. They succeed if

$$a \oplus b = xy, \tag{1}$$

where the plus operation $\oplus$ is modulo 2, otherwise they fail.

The best classical strategy succeeds with a probability of $3/4$. This can be achieved by, e.g., both Alice and Bob always outputting 0. It can be shown that no classical strategy performs better than this simple strategy.

Next, we define two closely related concepts: the CHSH inequality and the CHSH test. The CHSH inequality basically states that the winning probability of the CHSH game is less than or equal to $3/4$,

$$\Pr(\mathrm{CHSH}) \leqslant 3/4. \tag{2}$$

Since $3/4$ is the success probability of the best classical strategy, a violation of the CHSH inequality indicates nonclassicality, or quantumness. A CHSH test is the procedure of testing a violation of the CHSH inequality.

Finally, note that the best quantum strategy can achieve a winning probability of $(\sqrt{2}+2)/4 > 3/4$ with Alice and Bob sharing an Einstein–Podolsky–Rosen (EPR) pair. For example,

*xma@tsinghua.edu.cn

 ©2016 American Physical Society

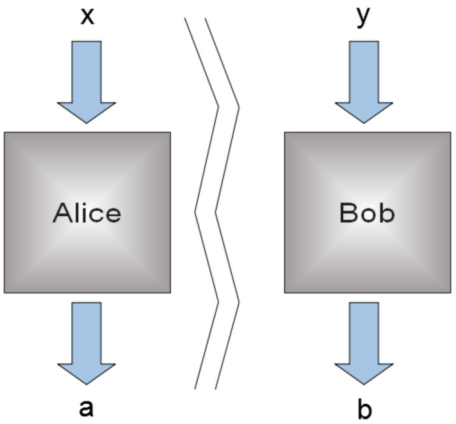

FIG. 1. CHSH game: Alice and Bob agree on a strategy before the game. During the game, they cannot communicate. Alice (Bob) is given an input bit x (y) and is required to output a bit a (b). Their goal is to let the outputs satisfy $a \oplus b = xy$.

Alice can choose the measurement bases σ_z, σ_x for the input $x = 0,1$ respectively and Bob can choose the measurement bases $(\sigma_z + \sigma_x)/\sqrt{2}$, $(\sigma_z - \sigma_x)/\sqrt{2}$ for the input $y = 0,\ 1$, respectively. Here σ_x and σ_z are Pauli matrices. During the game, after Alice (Bob) receives x (y), she (he) measures in the corresponding basis and outputs the binary measurement outcome. It can be calculated that this strategy achieves a winning probability of $(\sqrt{2}+2)/4$, which is larger than $3/4$.

B. Protocol description

The DIQKD protocol proceeds as follows: Faithful devices of Alice and Bob should share EPR pairs before the measurement. Keys are generated by Alice and Bob each measuring the qubits in the computational basis. To ensure that an EPR pair is indeed shared, Alice and Bob occasionally perform the CHSH test to check if the CHSH inequality is violated.

In the protocol, Alice and Bob receive inputs $x \in \{0,1,2\}$ and $y \in \{0,1\}$, respectively. The inputs $x,y \in \{0,1\}$ are used to test the CHSH inequality. If it works ideally, then $P(a \oplus b = xy) = \cos^2(\pi/8)$ for all inputs $x,y \in \{0,1\}$. When Alice's input is $x = 2$, Alice measures her part in the same basis as Bob measures when $y = 1$. Clearly, in the case $(x,y) = (2,1)$, Alice and Bob get identical keys when they share $(|00\rangle + |11\rangle)/\sqrt{2}$. The complete protocol is outlined in Table I.

C. Vazirani–Vidick security proof

Due to disturbance in the channel and/or eavesdropping, the raw keys are usually not identical or secure. Thus one needs to perform error correction to make the keys identical and to perform privacy amplification to extract the secure part of the keys. The secure part of the keys can be quantified by the smooth min entropy conditioned on Eve's information. After error correction and privacy amplification, the main term of the key rate is

$$R = H_{\min}^{\epsilon}(K|E) - I_{\mathrm{ec}}, \tag{3}$$

where $H_{\min}^{\epsilon}(K|E)$ is the smooth min entropy of the raw key and I_{ec} is the error correction (i.e., information reconciliation) cost.

In the DIQKD protocol, the information reconciliation cost is $I_{\mathrm{ec}} = h(e)$ [4], where $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the binary Shannon entropy. Note that this bound is the Shannon limit. In practice, the reconciliation cost is often taken to be $h(e)$ times a multiplicative factor for efficient computation [19,20]. For simplicity, we will omit this multiplicative factor in this work.

Asymptotically, the smooth min entropy part is

$$H_{\min}^{\epsilon}(K|E) = -\frac{177}{50}\log_2\left(\frac{11}{12} + 0.4986\sqrt{\delta}\right). \tag{4}$$

Thus Vazirani–Vidick key-rate formula is [21]

$$R = -\frac{177}{50}\log_2\left(\frac{11}{12} + 0.4986\sqrt{\delta}\right) - h(e). \tag{5}$$

We will take this as the baseline and later give an improved version.

III. IMPROVED KEY RATE

To improve the key rate, we examine the Vazirani–Vidick security proof in detail and optimize the parameters.

The security proof consists of three steps. The first step is to show if an adversary can predict all outcomes with a small probability, then there exists a run whose outcome he can accurately predict with a high probability conditional on outcomes of previous runs. This is done by quantum reconstruction paradigm [22]. It asserts that there exists an advice string Z of length $H_{\min}(A|E)$ (where A is the raw output and E is the adversary), such that, given Z, Eve can predict A correctly with a non-negligible probability.

The second step is to show that Eve's ability to predict one outcome with a high probability leads to signaling between the devices, contradicting the nonsignaling assumption. This fact

TABLE I. The devices of Alice and Bob are supposed to share n EPR pairs before the protocol. At the end of the protocol, if it does not abort, Alice and Bob share a key of length Rn, where R is the key rate given by Eq. (3).

1. **Main procedure:** For each round $i \in \{1,\ldots,n\}$ in n rounds, the devices of Alice and Bob randomly receive $x \in \{0,1,2\}$ and $y \in \{0,1\}$ and output $a,b \in \{0,1\}$, respectively.
2. **Testing:** Set noise parameters e, δ. A random subset of $\{1,\ldots,n\}$ is chosen as the set of tests, as illustrated in Fig. 2. Alice and Bob announce the outputs of these tests. For tests with inputs $x,y \in \{0,1\}$, they record the fraction of winning the CHSH game as $\cos^2(\pi/8) - \delta$. For tests with inputs $(x,y) = (2,1)$, they record the fraction of identical outputs as $1 - e$.
3. **Extraction:** The non-test rounds with inputs $(x,y) = (2,1)$ are used to generate a raw key. Finally, Alice and Bob perform standard error correction and privacy amplification on the raw key to obtain a secret identical key of length Rn. If $R \leqslant 0$, the protocol aborts.

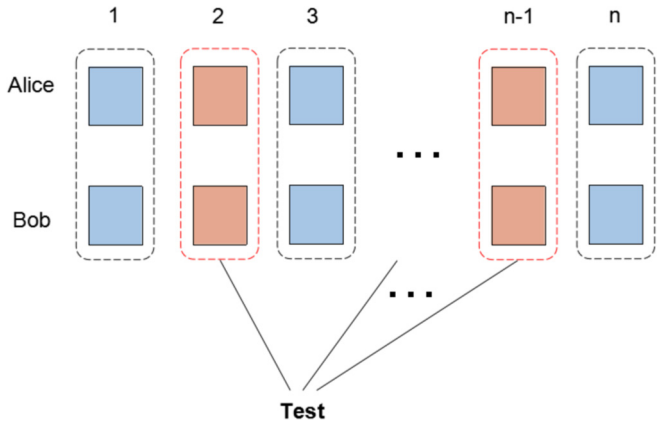

FIG. 2. Illustration of the protocol. Alice and Bob measure their joint states in n rounds, among which a red subset is randomly selected for tests.

is intuitive and can be quantitatively proved through a guessing game.

The third step is to deal with the conditioning and to show that it has almost no effects, because the conditioning on previous outcomes is not taken into consideration in the second step. The key idea to solve the third step is that, if Alice and Bob's shared state were dependent on Eve's input, they could have made a measurement on this postmeasurement state to recover Eve's input, contradicting the nonsignaling condition.

Combining these three steps, one can get a lower bound on the asymptotic smooth min entropy [21],

$$H_{\min}^{\epsilon}(K|E) = -6(1-\tau')\log_2\left(\frac{11}{12} + \frac{3\sqrt{\delta}}{2^{13/4}\sqrt{1-\tau}}\right), \quad (6)$$

where $\tau + \tau' > 1$. Different τ and τ' correspond to different smooth min entropy.

Taking $\tau = 60/100$, $\tau' = 41/100$ recovers Eq. (4). However, this choice does not necessarily give the optimal key rate and optimizing τ and τ' may improve the key rate. Thus we obtain a new estimate on the smooth min entropy by performing a numerical optimization on Eq. (6) with the constraint $\tau + \tau' > 1$,

$$H_{\min}^{\epsilon}(K|E) = \max_{\tau+\tau'>1} -6(1-\tau')\log_2\left(\frac{11}{12} + \frac{3\sqrt{\delta}}{2^{13/4}\sqrt{1-\tau}}\right). \quad (7)$$

By making the approximations $\ln(1-x) = -x$ and $\tau + \tau' = 1$, an analytical form of optimal τ can be derived in the infinite-key-size limit,

$$\tau_{\text{opt}} = 1 - \left(\frac{\delta^{-1/6}}{c} - \frac{2c\delta^{1/6}}{9} - 2^{-\frac{9}{4}}\delta^{\frac{1}{2}}\right)^{-2}, \quad (8)$$

where $c = 11^{-1/3} \times 3 \times 2^{-1/12}$. The details of derivation are in Appendix A.

For the other parameter τ', since the constraint is $\tau + \tau' > 1$, τ' can be chosen arbitrarily close to $1-\tau$ in the infinite-key-size limit. For the finite-key-size case, in order to optimize the key rate, τ' should not be chosen infinitely close to $1-\tau$ due to statistical fluctuation and a small gap is necessary. In all following simulations, τ' will be taken to be $1-\tau+0.01$. Note that this gap 0.01 actually makes almost negligible change to the key rate compared to the case of no gap. We leave it as an interesting open question to explore how this gap affects the finite-key analysis.

Figure 3 shows the numerically optimal τ and the approximately optimal τ_{opt} corresponding to different noise levels. For each noise level, the numerically optimal τ is obtained by sweeping its range $[0,1]$ at a step of 0.001 and selecting the value of τ that achieves the highest smooth min entropy (7). The original value of τ taken in the Vazirani–Vidick proof is also shown as the dot-dashed line in Fig. 3 for comparison. It can be seen that the best τ and τ', instead of constants as taken in the Vazirani–Vidick proof, are actually very sensitive to the noise level. When the noise parameter e is small, the best τ and τ' have a large gap of 1, and when e is the maximum allowable, the best τ and τ' is about equal. Figure 3 also shows that our analytical τ_{opt} is very close to the best τ, in contrast to the big gap between the best τ and the τ taken in the Vazirani–Vidick proof. This difference will be reflected in the comparison of key rates later.

Plugging the approximately optimal τ_{opt} into Eq. (7), we obtain the new analytical key-rate formula

$$R = -h(e) - 6\left[1 - \left(\frac{\delta^{-1/6}}{c} - \frac{2c\delta^{1/6}}{9} - 2^{-\frac{9}{4}}\delta^{\frac{1}{2}}\right)^{-2}\right] \times \log_2\left(\frac{11}{12} + \frac{3\left(\delta^{\frac{1}{3}}/c - 2c\delta^{\frac{2}{3}}/9 - 2^{-\frac{9}{4}}\delta\right)}{2^{13/4}}\right), \quad (9)$$

where $c = 11^{-1/3} \times 3 \times 2^{-1/12}$. By taking $e = 0$ and $R = 0$ in the key-rate formula, it can be calculated that the maximal tolerable δ is 0.0698. This is smaller than the difference between the classical and quantum winning probability, which is $\cos^2(\pi/8) - 3/4 = 0.1036$. It is an interesting question to close this gap.

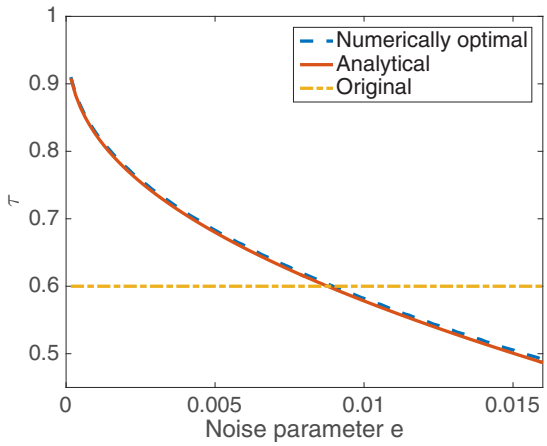

FIG. 3. The solid line shows the value of τ corresponding to the optimized key rate. It decreases from 1 to 0.49 as e increases from 0 to 1.6%. The dashed line shows the analytical approximation τ_{opt}, which is very close to the numerically optimal τ. The dot-dashed line stands for the value of τ in the Vazirani–Vidick proof.

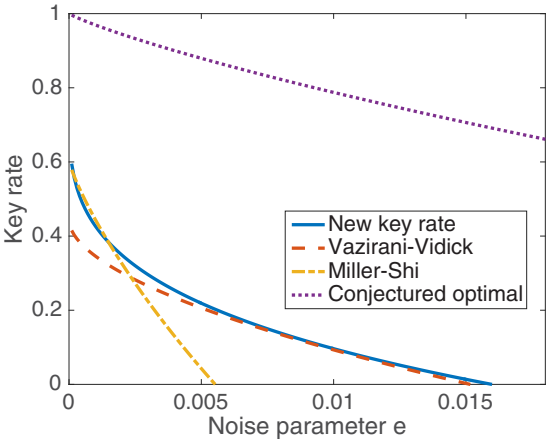

FIG. 4. The key rate of our optimized formula, the Vazirani–Vidick proof, the Miller–Shi proof, and the conjectured optimal key rate. For every value of noise parameter e, our key rate (solid line) is better than the Vazirani–Vidick proof (dashed line). The maximum allowable noise is also increased from 1.5% to 1.6%. Comparing with the Miller–Shi proof (dot dashed line), our work also has superior performance in most parameter regions.

To further evaluate the performance of our key-rate formula, we plot the relation between the key rate and the noise parameter and compare with the baseline. For simplicity, we take $\delta = e$ in the plot and all simulations afterwards. This should not be confused with the fact that δ and e are generally different.

Figure 4 shows that the new key rate is always larger than the baseline (Vazirani–Vidick) and has better noise tolerability. The maximum error tolerability is around 1.6%, which will be used in the subsequent section. We also compare our result with the Miller–Shi proof [17] (see Appendix B for details) and simulations show that our key rate is also higher in most parameter regions.

As a natural extension, we ask whether we can further improve the key rate. Although we do not attempt to answer the question directly here, inspired by the Shor–Preskill proof [3], we propose a conjectured optimal key rate and plot it as the dotted line in Fig. 4 (see Appendix C for details). It can be seen that our analytic key rate is not too far from this conjectured optimal key rate.

IV. EXPERIMENT CRITERIA

Given noise parameters e, δ, to ensure that the protocol does not always abort, one should put some requirements on the experimental instruments. In this section, we examine the constraints on the actual experimental equipment parameters and pave the way towards an experimental realization of DIQKD.

A. Parameters of experimental device

In the following, we mostly focus on examining the measurement. This is because the source, which is an EPR pair, could be prepared almost perfectly, while the main imperfections lie in the measurement. In addition, the misalignment of the EPR pair source can be transferred to the measurement misalignment.

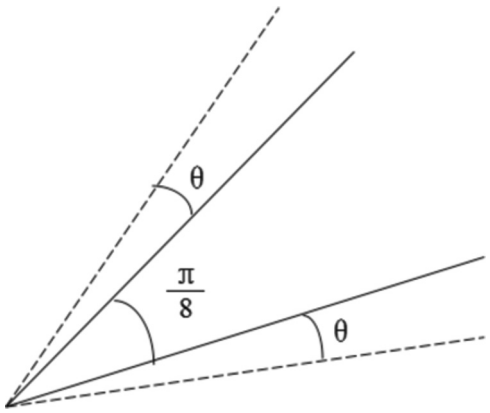

FIG. 5. Deviation of angle. The two solid lines stand for the ideal basis direction, while the two dashed lines stand for the deviated basis direction. The angle between the two deviated basis direction is shown to be $\pi/8 + 2\theta$.

In practical implementations, we allow the device to output a loss (i.e., there is no measurement outcome). In this case, we further add a step that when Alice (Bob) has a loss, she (he) assigns a random bit as the outcome. We denote the total transmittance (the probability of a photon to be detected, including channel loss and detector efficiency) as η, the detector misalignment angle as θ, which is the deviation from the basis that it should perform the measurement, and the dark count per detector as Y_0. Note that the misalignment is sometimes characterized by the misalignment error e_d in the QKD literature [20], which is related to θ by $e_d = \sin^2\theta$.

When Alice and Bob measure in the same basis, the probability of getting identical outputs should be larger than $1 - e$. According to the protocol, when playing the CHSH game, the success probability should be larger than $\cos^2(\pi/8) - \delta$. We now examine these two conditions and begin with the CHSH game.

If the only imperfection is that the transmittance is not 1, then the winning probability of the CHSH game becomes $\eta^2[\cos^2(\pi/8) - 1/2] + 1/2$ because, when both detectors react, they could win the game with probability $\cos^2(\pi/8)$ and otherwise they win with probability $1/2$.

If the only imperfection is detector misalignment, then the winning probability becomes $\cos^2(\pi/8 + 2\theta)$. The reason is as follows: Recall that the probability of winning is the inner product of two bases whose angle is ideally $\pi/8$ [18]. Since the basis of both of Alice and Bob might deviate by at most θ, the angle between them will be at most $\pi/8 + 2\theta$. An illustration is shown in Fig. 5.

If the only imperfection is dark count, then the violation becomes $\cos^2(\pi/8) - 2Y_0$, because either detector receiving a dark count is regarded as a failure. In all, with all imperfections, the violation is $\eta^2[\cos^2(\pi/8 + 2\theta) - \frac{1}{2}] + \frac{1}{2} - 2Y_0$.

Similarly there is a parallel requirement for getting identical outputs when measuring in the same basis. Considering the imperfection separately, the probability of getting identical outputs becomes $\eta^2/2 + 1/2$ instead of 1 for efficiency, $\cos^2(2\theta)$ for misalignment, and $1 - 2Y_0$ for dark count. Thus,

with all imperfections, it becomes $\eta^2[\cos^2(2\theta) - \frac{1}{2}] + \frac{1}{2} - 2Y_0$.

Thus, we have the following two constraints on the experimental parameters for given noise parameters e and δ:

$$\eta^2\left[\cos^2\left(\frac{\pi}{8}+2\theta\right)-\frac{1}{2}\right]+\frac{1}{2}-2Y_0 \geqslant \cos^2\frac{\pi}{8}-\delta,$$
$$\eta^2\left(\cos^2(2\theta)-\frac{1}{2}\right)+\frac{1}{2}-2Y_0 \geqslant 1-e. \tag{10}$$

In fact, when $\delta = e$, the second inequality can induce the first inequality. Thus the combined constraints of the two inequalities will be equivalent to the single constraint of the second inequality in our simulation. We will refer to this constraint as *the basic constraint*.

B. Refined rate with loss position

We can improve the error correction cost by a simple technique developed in Ref. [23]. Note that Alice can use her knowledge on which positions are losses when performing error correction. By splitting the error e into the loss part which has error $1/2$ and the no-loss part which has error e', the error-correction cost decreases from $I_{\text{ec}} = h(e)$ to

$$I_{\text{ec}} = 1 - \eta + \eta h(e'), \tag{11}$$

where $e' = (2e - 1 + \eta)/(2\eta)$. We later refer to this constraint as *the refined constraint*.

C. Simulation

Next we examine the allowable experimental-parameter region. The maximum allowable noise parameter e is taken to be 1.6% according to the previous section. Since Y_0 is usually small in practical systems, typically on the order of 10^{-6}, it can be omitted. We plot the relation between the misalignment θ and the total transmittance η in Fig. 6. From Fig. 6, the detector efficiency for the basic constraint needs to be at least 98.4%, and for the refined constraint at least 97.3%. They are both much higher than the requirement for the violation of CHSH inequality, which is 67%.

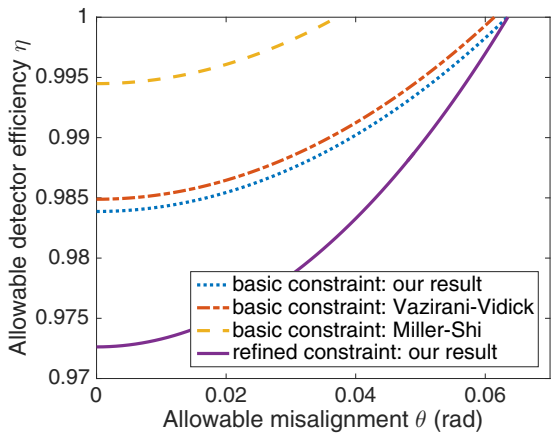

FIG. 6. Transmittance η vs misalignment θ. As more misalignment is allowed, the required minimum detector efficiency becomes higher. Four cases are considered, basic constraint for our result, Vazirani–Vidick, Miller–Shi, and refined constraint for our result.

D. Asymmetric case

Usually Alice's and Bob's experimental instruments are not identical. Thus their experimental parameters may differ. Furthermore, if one instrument has very high quality, the requirement on the other instrument can be loosened. In other words, there are some trade-offs between the parameters of the two instruments held by Alice and Bob. We denote the efficiencies of the two measurement systems as η_1, η_2, and the misalignments as θ_1, θ_2, respectively. Then following the previous calculations, an equivalent average efficiency and misalignment have the forms

$$\theta = \frac{\theta_1 + \theta_2}{2}, \tag{12}$$

$$\eta = \sqrt{\eta_1 \eta_2}. \tag{13}$$

This in particular implies that, if one detector has perfect unity detector efficiency, then the other detector only needs an efficiency of 95%.

E. Visibility

Finally, the single photon EPR source in the above analysis can be replaced by a more practical parametric down conversion (PDC) source. The fidelity of the photon source is usually characterized by the visibility. The multiphoton components should usually be reduced because they have lower visibilities [24]. Denote the overall visibility as $\mathcal{V}$. When Alice and Bob measure with the same orthogonal basis, e.g., $\{H, V\}$, they will get four results, H and H, H and V, V and H, and V and V. The visibility is defined as

$$\mathcal{V} = \frac{P_{HH} + P_{VV} - P_{HV} - P_{VH}}{P_{HV} + P_{VH} + P_{HH} + P_{VV}}, \tag{14}$$

where P_{HH} for instance is the probability of getting the results H and H for the two parties. The results H and V, V and H contribute to the error rate e, thus

$$e \geqslant \frac{P_{HV} + P_{VH}}{P_{HV} + P_{VH} + P_{HH} + P_{VV}}. \tag{15}$$

Comparing Eq. (14) and Eq. (15), the visibility should satisfy $\mathcal{V} \geqslant 1 - 2e$. Substituting the constraint on the error rate $e \leqslant 1.6\%$, we get that the visibility of the source in DIQKD must be higher than 96.8%.

V. CONCLUSION

In summary, we have explicitly put the requirements on the experimental devices and in the meantime improved the key rate for DIQKD. There are a few interesting further works that are worth investigating. On the experimental side, this includes selecting the appropriate intensity for the PDC source to maximize the key rate. Some assumptions on the detectors might need to be employed in the near future because the current requirement of detector efficiency for a strict demonstration of DIQKD is much higher than most experimental systems.

On the theoretical side, first it is interesting to explore the optimal key rate for the DIQKD protocol, as already mentioned previously. This might be hard because similar questions are

still wide open even for trusted device QKD protocols, such as the well-known BB84 protocol. Second, lowering the detector efficiency requirement, or better still, separating the detector efficiency from the protocol is of vital importance for practical realizations of DIQKD. A possible solution is to replace CHSH inequality used in the protocol by other Bell inequalities. Third, for a loss, it is interesting to consider employing a fixed assignment instead of a random assignment and examine which has better performance. In this case the key rate may need to be rederived because a fixed assignment leads to a decrease in min entropy of the raw key.

ACKNOWLEDGMENTS

We thank C. Miller, Y. Shi, and T. Vidick for useful discussions. This work was supported by the 1000 Youth Fellowship program in China.

APPENDIX A: PROOF OF EQ. (8)

By substituting $\tau' = 1-\tau$ and taking the derivative over τ on Eq. (7), we have

$$\ln\left(\frac{11}{12}+\frac{3\sqrt{\delta}}{2^{13/4}\sqrt{1-\tau}}\right)+\tau\frac{\frac{3\sqrt{\delta}}{2^{13/4}(1-\tau)^{3/2}}\frac{1}{2}}{\frac{11}{12}+\frac{3\sqrt{\delta}}{2^{13/4}\sqrt{1-\tau}}}=0. \quad \text{(A1)}$$

By making the approximation $\ln(1-u)=-u$ and letting $x=\sqrt{1-\tau}$, we obtain a cubic equation

$$\frac{\sqrt{\delta}c_1x^3}{2}+c_1^2\delta x^2+\frac{\sqrt{\delta}c_1x}{3}-\frac{11}{144}=0, \quad \text{(A2)}$$

where $c_1=3\times 2^{-13/4}$. Solving this equation and taking leading terms yields Eq. (8).

APPENDIX B: KEY RATE OF MILLER AND SHI

In this Appendix, we review the key-rate formula of Miller and Shi [17]. The Miller–Shi proof actually holds for a more general class of Bell games, the so-called quantum XOR game. For fair comparison, we take the CHSH game that involves two parties because other quantum XOR games require at least three space-separated parties and are more resource demanding.

The key rate used in our simulation is given by

$$R=h\left(w_G-\frac{1}{2}\right)-2h\left(\frac{e}{v_G}\right), \quad \text{(B1)}$$

where the winning probability w_G is $(2+\sqrt{2})/4$ and the trust coefficient v_G is 0.0732 for CHSH game. (In Sec. B 1, we explicitly calculate this v_G.) The noise parameter e is defined in the main text and we note that they originally use the symbol η.

Actually, in a subsequent paper of Miller and Shi [25], they derive another bound on the privacy amplification part for arbitrary nonlocal games. Combining it with the reasoning of their first work would yield a key rate of

$$r=h\left(w_G-\frac{1}{2}\right)+\frac{2\left(\frac{\sqrt{2}-1}{4}-e\right)^2}{3\ln 2}-1 \quad \text{(B2)}$$

for the CHSH game. However, this key rate is always negative and we thus omit it from our simulation.

1. Calculation of v_G for Clauser–Horne–Shimony–Holt game

First we present the definition of v_G and then calculate it for the CHSH game. Mathematically, if there exists some N whose square is the identity matrix and that anticommutes with $\sigma_x\otimes I$ satisfies $||M-v_GN||\leqslant q_G-v_G$ where M will be defined below and q_G is $2w_G-1=\sqrt{2}/2$ for the CHSH game, then v_G is the trust coefficient of the game.

We perform the calculation following the Greenberger–Horne–Zeilinger (GHZ) game example in Sec. I.3 of Ref. [17]. The polynomial of CHSH is

$$P(\zeta_1,\zeta_2)=\tfrac{1}{4}(1+\zeta_1+\zeta_2-\zeta_1\zeta_2), \quad \text{(B3)}$$

where ζ_1,ζ_2 are complex number of norm 1. The matrix M is defined as

$$M=\begin{pmatrix} & & P(\xi_1,\xi_2)\\ & P(\bar{\xi_1},\xi_2) & \\ P(\bar{\xi_1},\bar{\xi_2}) & & \end{pmatrix}, \quad \text{(B4)}$$

where $\mathrm{Im}(\xi_1)\geqslant 0$, $\mathrm{Im}(\xi_2)\geqslant 0$.

Take

$$N=-\sigma_y\otimes\mathbb{I}_2=\begin{pmatrix} & & i\\ & -i & \\ -i & & \end{pmatrix}, \quad \text{(B5)}$$

which clearly anticommutes with $\sigma_x\otimes\mathbb{I}_2$ and satisfies $N^2=\mathbb{I}_4$. We need to find a v_G such that

$$||M-v_GN||\leqslant q_G-v_G. \quad \text{(B6)}$$

The answer is $v_G=(\sqrt{2}-1)/4=0.104$. To prove this, we just need to show that every reverse diagonal term of $M-v_GN$ is bounded by q_G-v_G.

(1) First we prove $|P(\xi_1,\bar{\xi_2})-v_Gi|\leqslant q_G-v_G$. Denote $\zeta_j=\cos\theta_j+i\sin\theta_j$, $j=1,2$. By direct calculation, we can show

$$|P(\zeta_1,\zeta_2)|=\sqrt{1+\sin\theta_1\sin\theta_2}/2. \quad \text{(B7)}$$

Since $\sin\theta_1\geqslant 0$ for ξ_1 [because $\mathrm{Im}(\xi_1)\geqslant 0$] and $\sin\theta_2\leqslant 0$ for ξ_2 [because $\mathrm{Im}(\bar{\xi_2})\leqslant 0$], thus $|P(\xi_1,\bar{\xi_2})|\leqslant 1/2$. Thus,

$$\begin{aligned}|P(\xi_1,\bar{\xi_2})-v_Gi|&\leqslant|P(\xi_1,\bar{\xi_2})|+|v_G|\\&\leqslant 1/2+v_G=1/2+(\sqrt{2}-1)/4=q_G-v_G.\end{aligned} \quad \text{(B8)}$$

Similarly $|P(\bar{\xi_1},\xi_2)+v_Gi|\leqslant q_G-v_G$.

(2) Next we prove $|P(\xi_1,\xi_2)-v_Gi|\leqslant q_G-v_G$. Actually, we can split $P(\zeta_1,\zeta_2)$ into two parts:

(i) $(\zeta_1+\zeta_2)/4$, which has length $\cos[(\theta_1-\theta_2)/2]/2$ and has angle $(\theta_1+\theta_2)/2$,

(ii) $(1-\zeta_1\zeta_2)/4$, which has length $\sin[(\theta_1+\theta_2)/2]/2$ and has angle $(\theta_1+\theta_2)/2-\pi/2$.

Since $0\leqslant\theta_1,\theta_2\leqslant\pi$ for ξ_1 and ξ_2, we divide $\theta_1+\theta_2$ into two cases: $[0,\pi]$ and $[\pi,2\pi]$. We can easily prove for each case $|P(\xi_1,\xi_2)-v_Gi|\leqslant q_G-v_G$ by noting the fact that $\cos[(\theta_1-\theta_2)/2]/2\leqslant 1/2$ and $\sin[(\theta_1+\theta_2)/2]/2\leqslant 1/2$.

For example, in the second case, which is the harder one, denote the angle between v_Gi and $(1-\zeta_1\zeta_2)/4$ as ϕ. Since

$\theta_1+\theta_2\in[\pi,2\pi]$, we have $\phi\in[0,\pi/2]$. Let $a=\cos[(\theta_1-\theta_2)/2]/2$, $b=\sin[(\theta_1+\theta_2)/2]/2$, we just need to show

$$(a-v_G\sin\phi)^2+(b-v_G\cos\phi)^2\leqslant(\sqrt{2}/2-v_G)^2. \quad \text{(B9)}$$

We expand its left-hand side (LHS) and perform a series of relaxations:

$$\begin{aligned}\text{LHS}&=a^2-2av_G\sin\phi+b^2-2bv_G\cos\phi+v_G^2\\&\leqslant a/2-2av_G\sin\phi+b/2-2bv_G\cos\phi+v_G^2\\&=a\left(\frac{1}{2}-2v_G\sin\phi\right)+b\left(\frac{1}{2}-2v_G\cos\phi\right)+v_G^2\\&\leqslant\left(\frac{1}{2}-2v_G\sin\phi\right)\Big/2+\left(\frac{1}{2}-2v_G\cos\phi\right)\Big/2+v_G^2\\&=(\sqrt{2}/2-v_G)^2,\end{aligned} \quad \text{(B10)}$$

which proves Eq. (B9).

Similarly, we can prove $|P(\bar{\xi}_1,\bar{\xi}_2)+v_Gi|\leqslant q_G-v_G$. This completes the proof of Eq. (B6).

APPENDIX C: OPTIMAL-KEY-RATE CONJECTURE

By the Shor–Preskill proof [3], the key rate is given by $R=1-h(e_{\text{bit}})-h(e_{\text{ph}})$. The term $h(e_{\text{bit}})$ is for error correction cost depending on the bit error rate e_{bit}. This bit error rate e_{bit} is the same as e in Eqs. (5) and (9).

The other term $h(e_{\text{ph}})$ is for privacy amplification depending on the phase error rate e_{ph}. Denote the deviation from the maximal winning probability of CHSH game as δ. We next bound e_{ph} by δ.

In a previous work on measurement-device independence [26], it was shown that an untrusted measurement with two outcomes can be restricted to a two-dimensional projective measurement without loss of generality. In light of this and by the fact that each party in the CHSH game has two inputs and two outputs, operating each untrusted device in a two-dimensional space seems to already include the worst adversarial case of device independence. Thus, it is reasonable to assume that both devices hold qubits and perform trusted qubit projective measurements.

Furthermore, we assume that the deviation from the maximal violation has two separate causes, the bit error and the phase error. Having a phase error e_{ph} and a bit error e is equivalent to the shared state between Alice and Bob being $|00\rangle-|11\rangle$ with probability e_{ph}, $|01\rangle+|10\rangle$ with probability e, and $|00\rangle+|11\rangle$ with probability $1-e_{\text{ph}}-e$. It can be calculated that both $|00\rangle-|11\rangle$ and $|01\rangle+|10\rangle$ will result in a success probability of $1/2$ for the CHSH game. Thus

$$\frac{2+\sqrt{2}}{4}(1-e_{\text{ph}}-e)+\frac{1}{2}e_{\text{ph}}+\frac{1}{2}e=\frac{2+\sqrt{2}}{4}-\delta. \quad \text{(C1)}$$

Rearranging the terms, we get $e_{\text{ph}}=2\sqrt{2}\delta-e$.

Thus our conjectured key rate is

$$R=1-h(e)-h(2\sqrt{2}\delta-e). \quad \text{(C2)}$$

Note that Eve can perform an attack that saturates this key-rate bound; namely, by preparing $|00\rangle-|11\rangle$ with probability $2\sqrt{2}\delta-e$, $|01\rangle+|10\rangle$ with probability e, and $|00\rangle+|11\rangle$ with probability $1-2\sqrt{2}\delta$.

[1] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (IEEE Press, New York, 1984), p. 175.
[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] C. E. Shannon, Bell Syst. Tech. J. **28**, 656 (1949).
[5] M. N. Wegman and J. L. Carter, J. Comput. Syst. Sci. **22**, 265 (1981).
[6] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, Quantum Inf. Comput. **4**, 325 (2004).
[7] B. Huttner, N. Imoto, N. Gisin, and T. Mor, Phys. Rev. A **51**, 1863 (1995).
[8] B. Qi, C.-H. F. Fung, H.-K. Lo, and X. Ma, Quantum Inf. Comput. **7**, 73 (2007).
[9] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Nat. Photonics **4**, 686 (2010).
[10] D. Mayers and A. Yao, in *Proceedings of 39th Annual Symposium on Foundations of Computer Science, 1998* (IEEE, New York, 1998), p. 503.
[11] A. Acin, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).
[12] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[13] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 010503 (2005).
[14] S. Pironio, A. Acin, N. Brunner, N. Gisin, S. Massar, and V. Scarani, New J. Phys. **11**, 045021 (2009).
[15] U. Vazirani and T. Vidick, Phys. Rev. Lett. **113**, 140501 (2014).
[16] B. M. Terhal, IBM J. Res. Dev. **48**, 71 (2004).
[17] C. A. Miller and Y. Shi, arXiv:1402.0489.
[18] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[19] G. Brassard and L. Salvail, in *Advances in Cryptology–EUROCRYPT'93* (Springer, Berlin, Heidelberg, 1993), p. 410.
[20] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, Phys. Rev. A **72**, 012326 (2005).
[21] U. Vazirani and T. Vidick, Phys. Rev. Lett. **116**, 089901(E) (2016).
[22] U. Vazirani and T. Vidick, in *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing* (ACM, New York, 2012), p. 61.
[23] X. Ma and N. Tkenhaus, Quantum Inf. Comput. **12**, 203 (2012).
[24] J. F. Hodelin, G. Khoury, and D. Bouwmeester, Phys. Rev. A **74**, 013802 (2006).
[25] C. A. Miller and Y. Shi, arXiv:1411.6608.
[26] Z. Cao, H. Zhou, and X. Ma, New J. Phys. **17**, 125011 (2015).