Biased decoy-state measurement-device-independent quantum key distribution with finite resources

Chun Zhou,[1,2] Wan-Su Bao,[1,2,*] Hai-long Zhang,[1,2] Hong-Wei Li,[1,2,3] Yang Wang,[1,2] Yuan Li,[1,2] and Xiang Wang[1,2]
[1]*Zhengzhou Information Science and Technology Institute, Zhengzhou, 450004, China*
[2]*Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*
[3]*Key Laboratory of Quantum Information,University of Science and Technology of China, Hefei, 230026, China*

(Received 23 July 2014; published 17 February 2015)

Measurement-device-independent quantum key distribution (MDI-QKD) can remove all the side-channel attacks from imperfections in the detection side. However, finite-size resources undoubtedly influence its performance and the achievable finite secret key rates of MDI-QKD are typically lower than that of standard decoy-state QKD. In this paper, we introduce the efficient decoy-state method with biased basis choice into the finite-key analysis and propose a decoy-state protocol for MDI-QKD. By applying vacuum + weak decoy-state method, we analytically derive concise formulas for estimating the lower bound of single-photon yield and the upper bound of phase error rate in the case of finite resources. The simulations show that proper basis choice combined with deliberate intensity choice can substantially enhance the performance of decoy-state MDI-QKD and, without a full optimization program, our protocol can bring a long-distance implementation (168 km on standard optical fiber) of MDI-QKD with a reasonable data size of total transmitting signals ($N = 10^{15}$).

DOI: 10.1103/PhysRevA.91.022313 PACS number(s): 03.67.Dd, 03.67.Hk

I. INTRODUCTION

Quantum key distribution (QKD), stemming from the first pioneering proposal presented in 1984 (traditionally known as the BB84 protocol) [1], allows two remote parties to acquire a secret key based on quantum mechanics. Since then, QKD has received considerable attention and great progress has been achieved in both theory and practice [2]. However, in practical situations, in a real QKD system there undoubtedly exist sorts of imperfections such as inefficient authentication of classical communication, finite-size effect of processing data, and drawbacks of significant setup. For the first two kinds of imperfections, a theoretical model can be derived to characterize their influences and some efforts have been made to tackle these problems [3–6]. For the imperfections occurring in the setup, some security loopholes may exist and quantum hacking strategies can be smartly designed to enable the eavesdropper to acquire the secret key of the practical QKD system [7–9]. One countermeasure, although difficult to implement, is trying to derive an efficient mathematical model to characterize the system fully and take all the side channels into account in security proof as comprehensively as possible [2]. And the other one, is trying to build the fully device-independent QKD (DI-QKD) system [10,11]. The security of DI-QKD can be guaranteed by violation of a Bell inequality [12] without knowing the detailed information of the practical setups. But, with present-day technologies, the DI-QKD is impractical and hard to realize sice it inherently requires a high detection efficiency to overcome the security loophole.

Measurement-device-independent QKD (MDI-QKD) scheme [13] is one of the approaches [13–16] that are intermediate between standard (device-dependent) QKD and DI-QKD. In MDI-QKD, Alice and Bob both send pulses to the detection system that can be considered as an oracle with input and output trusted. The oracle may be controlled even by an eavesdropper, say Eve. She typically performs a partial Bell-state measurement and announces the results to Alice and Bob for distilling a secret key. Since MDI-QKD's security is guaranteed by the entanglement swapping techniques [14] and does not rely on the detection party, it can remove all detector side channels. After the invention of MDI-QKD, it attracts wide concentration and has been studied from different aspects both in theory and experiment [17–23].

A single-photon source is preferable for the realization of MDI-QKD. But an ideal single-photon source is unreasonable and still commercially unavailable with current technology. Practical photon sources such as the weak coherent sources (WCS) and the spontaneous-parametric-down-conversion sources (SPDCS) can be used as the source for MDI-QKD. Hence, one should apply the decoy-state method [24–26] to make sure of its unconditional security. Luckily, many attempts were made at estimating the bound of key rates such as the situations for WCS [17,18,27], heralded single-photon source [28], practical SPDCS [29], and the general imperfect single-photon source [19,30]. Note that the finite-key analysis has been carried on meticulously in a deepgoing way for the typical standard QKD [31–38] and the recent one based on the generalized uncertainty relation [6] has also been extended to situations like one-sided device-independent QKD [39], passive decoy-state protocol [40], and B92 protocol [41]. However, for decoy-state MDI-QKD, the finite-key analysis by far [42–44] is insufficient, since the achievable finite secret key rates of the existing results are not high and close enough to the asymptotic case given the same experimental data. Thus, efforts have been made to upgrade the theoretical results so as to obtain a higher key rate of decoy-state MDI-QKD [45,46]. A concise security bound against general attacks for typical standard decoy-state QKD with finite resources was recently presented by Lim *et al.* [47] and that for MDI-QKD is extremely urgent to be further studied.

[*] 2010thzz@sina.com

In this work, inspired from the proposed efficient decoy-state BB84 protocol [48,49], we apply a tight finite-key analysis for the decoy-state MDI-QKD where the basis is chosen with a biased probability and the intensities of different types of state are selected properly to optimize the achievable finite secret key rate. Most importantly, by combining the analytical method presented in [30] with the finite-key analysis, the formulas for estimating the single-photon yield and error rate of decoy-state MDI-QKD are recalculated, which can be proven to be rigorously concise in the finite-key case. In addition, we introduce the bound for estimating the phase error rate [36] into the decoy-state MDI-QKD with finite resources. And it should be noted that our security analysis is based on the recent security proof technique [50]. Hence, we are able to give concise finite-key security bounds that are valid against the most general attacks. Furthermore, we compare the performance of our bound with that obtained in [44] given the same experimental data under a realistic fiber-based system model. The simulations show a long-distance implementation (168 km on standard optical fiber) of MDI-QKD is acceptable with a reasonable data size of total signals ($N = 10^{15}$) and verify that our analytical security bounds for decoy-state MDI-QKD are profoundly concise with practical postprocessing block sizes.

The paper is constituted as follows. In Sec. II, we propose an efficient decoy-state MDI-QKD protocol with biased basis choice and specify its execution steps. Section. III fixes the security notions and introduces the formalism for calculating the achievable finite secret key rates. The main results of this work, i.e., concise formulas of bounding the yield and bit error rate for the single-photon events, are shown in Sec. IV. Section V numerically simulates our bounds and Sec. VI concludes the paper.

II. PROTOCOL DECLARATION

In this section, we propose a biased decoy-state MDI-QKD protocol. The protocol uses one-decoy settings and phase-randomized laser pulses to guarantee its security in the state-preparation step. We consider the vacuum + weak decoy-state method in which both Alice and Bob randomly modulate states with the three intensities μ, υ, and 0 ($\mu > \upsilon > 0$), which are named as the signal state, decoy state, and vacuum state, respectively. For different choices of intensity in Alice and Bob's modulators, the bases X and Z are selected with different probabilities.

Note that the error rates in X and Z bases are intrinsically not symmetric for MDI-QKD [13,18,21]; then one can smartly set different bases for specific choices of intensity to maximize the finite secret key rate. If we consider the phase-encoded MDI-QKD given by Ma *et al.* [18], the key bit is encoded in time bin 0 or time bin 1 under the Z basis and encoded into the relative phases 0 or π between the two time bins by a phase modulator under the X basis. However, the multiphoton component in the transmitted phase-randomized pulses may cause accidental coincidence, which introduces a 1/2 bit error rate in the X basis [21]. Hence, this will result in a higher quantum bit error rate with the X basis than that with the Z basis. It is the same case for the original polarization-based MDI-QKD protocol [13] where the rectilinear basis serves as the Z basis and the diagonal basis acts as the X basis. Thus, in our protocol, the final secure key is extracted from the successful detection events when both Alice and Bob choose the Z basis. Most importantly, to reduce the cost sacrificed in the parameter estimation step, the states are prepared only in the Z basis if Alice and Bob both choose the signal intensities. When Alice and Bob choose other intensities, the states are prepared in the X basis and Z basis with probabilities of q_X and $1 - q_X$, respectively. In the following, we provide a detailed description of our protocol.

TABLE I. Basis choice strategy for Alice and Bob's different intensity modulations.

Alice/Bob	0	υ_b	μ_b
0	–	X or Z	X or Z
υ_a	X or Z	X or Z	X or Z
μ_a	X or Z	X or Z	Z

Preparation. For the ith pulses emitted from the lasers, Alice and Bob modulate the intensities $\alpha_i \in \mathcal{A} := \{\mu_a, \upsilon_a, 0\}$ and $\beta_i \in \mathcal{B} := \{\mu_b, \upsilon_b, 0\}$ with the probabilities $p_{a_i} \in \{p_{\mu_a}, p_{\upsilon_a}, 1 - p_{\mu_a} - p_{\upsilon_a}\}$ and $p_{b_i} \in \{p_{\mu_b}, p_{\upsilon_b}, 1 - p_{\mu_b} - p_{\upsilon_b}\}$, respectively. Then, they assign bit values chosen uniformly from the random generators to prepare the phase-randomized laser pulses under the basis $k_i \in \{X, Z\}$ following the choice strategy shown in Table I, where the X basis is chosen with probability q_X and the Z basis with $1 - q_X$. Alice and Bob then send their pulses to Charlie via the quantum channel.

Measurement and sifting. Charlie conducts the partial Bell state measurement for the received pulse, and informs Alice and Bob of his measurement result through a public channel. Alice and Bob then publicly compare their basis for each pulse via an authenticated classical channel. For the pulse when they use the same basis and Charlie announces a successful measurement event, either Alice or Bob applies bit flip, and record the bit value in $y_i' \in \{0,1\}$ as a raw key. Then Alice and Bob are able to gather the pulse satisfying the following sets: $\mathcal{N}_{00} := \{i : (\alpha_i = 0) \wedge (\beta_i = 0) \wedge (y_i' \neq \phi)\}$, $\mathcal{X}_{\alpha\beta} := \{i : (\alpha_i = \alpha) \wedge (\beta_i = \beta) \wedge (k_i = X) \wedge (y_i' \neq \phi)\}$ with $\alpha \in \mathcal{A}, \beta \in \mathcal{B}$ except $\alpha\beta = \mu_a\mu_b$ and $\alpha\beta = 00, \mathcal{Z}_{\alpha\beta} := \{i : (\alpha_i = \alpha) \wedge (\beta_i = \beta) \wedge (k_i = Z) \wedge (y_i' \neq \phi)\}$ with $\alpha \in \mathcal{A}$, $\beta \in \mathcal{B}$ except $\alpha\beta = 00$. Here, $y_i' \neq \phi$ means that the ith pulse brings a successful measurement event in the Charlie's side. Then they check for $|\mathcal{N}_{00}| \geqslant n_{00}$, $|\mathcal{X}_{\alpha\beta}| \geqslant n_{\alpha\beta}^X$ for $\alpha\beta \neq \mu_a\mu_b$ and $\alpha\beta \neq 00$, $|\mathcal{Z}_{\alpha\beta}| \geqslant n_{\alpha\beta}^Z$ for $\alpha\beta \neq 00$, where $n_{00}, n_{\alpha\beta}^X(\alpha\beta \neq \mu_a\mu_b,\ \alpha\beta \neq 00)$ and $n_{\alpha\beta}^Z(\alpha\beta \neq 00)$ are postselected for parameter estimation and raw key generation. They repeat the preparation and measurement processes until these conditions are all satisfied. And they record the total number of pulses sent from Alice and Bob as N.

Parameter estimation. For the pulses prepared under the Z basis, Alice and Bob randomly choose a set of size $n^Z = \sum_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} n_{\alpha\beta}^Z$ from $\mathcal{Z} = \cup_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} \mathcal{Z}_{\alpha\beta}$ as the raw key pair $(\mathbf{X}_A, \mathbf{X}_B)$, where $n_{00}^Z = n_{00}(1 - q_X)$. Here, n^Z is considered as the postprocessing block size and all intensity levels under the Z basis are used for raw key generation. Then, they announce a sample set of size $n^X = \sum_{\alpha\in\mathcal{A},\beta\in\mathcal{B},\alpha\beta\neq\mu_a\mu_b,\alpha\beta\neq 00} n_{\alpha\beta}^X$ from

$\mathcal{X} = \cup_{\alpha\in\mathcal{A},\beta\in\mathcal{B}}\mathcal{X}_{\alpha\beta}$ to estimate the single-photon detections and calculate the number of bit errors $m^X_{\upsilon_a\upsilon_b}$, $m^X_{0\upsilon_b}$, and $m^X_{\upsilon_a 0}$ in the X basis, respectively. Given $n^Z_{\alpha\beta}$, Alice and Bob are able to estimate the number of single-photon detections s^Z_{11} in $(\mathbf{X}_A,\mathbf{X}_B)$. Likely, they can estimate the single-photon detections s^X_{11} among the tested sample events in the X basis given $n^X_{\alpha\beta}$. Then, with the observations $m^X_{\upsilon_a\upsilon_b}$, $m^X_{0\upsilon_b}$, and $m^X_{\upsilon_a 0}$, they can compute the number of single-photon bit errors c^X_{11} under the X basis. The phase error rate, denoted as $\varphi_Z := c^Z_{11}/s^Z_{11}$, cannot be directly measured from the raw key bits. But it can be bounded with the tested bit error rate c^X_{11}/s^X_{11} from the sample events by random sampling theory. Finally, they check if the estimated phase error rate φ_Z satisfies the condition $\varphi_Z \leqslant \varphi_{\text{tol}}$, where the phase error bound φ_{tol} is predetermined. The protocol only aborts if $\varphi_X > \varphi_{\text{tol}}$, otherwise they proceed to the next step.

Postprocessing. First, an error correction algorithm is applied to the raw key pair $(\mathbf{X}_A,\mathbf{X}_B)$, which reveals at most λ_{EC} bits of information. Then, by using two-universal hash functions that consume $\lceil\log_2(1/\varepsilon_{\text{hash}})\rceil$ bits of information, Alice and Bob conduct an error-verification step to ensure that their corrected keys are identical. Here, $\varepsilon_{\text{hash}}$ denotes the probability that a pair of different keys passes the error-verification step. At last, according to the information that leaks to the eavesdropper, the privacy amplification step compresses their keys to some extent for extracting a secret key pair $(\mathbf{S}_A,\mathbf{S}_B)$ with length $\ell = |\mathbf{S}_A| = |\mathbf{S}_B|$.

It should be noted that our protocol is apparently different from that of [44]. First of all, in the state preparation step, the basis is chosen biased with the purpose of maximizing the finite secret key rate in our protocol, which consumes a lower number of pulses for parameter estimation than the one in [44]. Secondly, in the parameter estimation step, all intensity levels including the vacuum states, decoy states, and signal states in the Z basis contribute to the raw keys in our protocol, which brings larger raw keys than the one in [44]. Thus, these fundamental advantages of our protocol can help us to improve the performance of decoy-state MDI-QKD with finite resources.

III. SECURITY DEFINITION AND SECRET KEY LENGTH

After the above protocol, a secret key pair $(\mathbf{S}_A,\mathbf{S}_B)$ with length ℓ is obtained. It is necessary to guarantee the security of the final key with quantified secrecy. First of all, we shall clarify the security criteria, which lays the foundation of our analysis. In this paper, we employ the notion of composable security based on trace distance into our analysis, which is initially proposed by Renner [51].

Definition 1 (composable security definition). The key pair $(\mathbf{S}_A,\mathbf{S}_B)$ that outputs from the protocol is considered to be ε-secure if it is both ε_{cor}-correct and ε_{sec}-secret. ε_{cor}-correct is satisfied only if $\Pr(\mathbf{S}_A \neq \mathbf{S}_B) \leqslant \varepsilon_{\text{cor}}$, i.e., the probability of $\mathbf{S}_A \neq \mathbf{S}_B$ will not exceed ε_{cor}. ε_{sec}-secret is satisfied only if $\frac{p_{\text{pass}}}{2}\|\rho_{SE} - U_S\otimes\rho_E\|_1 \leqslant \varepsilon_{\text{sec}}$, where S represents either of the keys $\mathbf{S}_A$ and $\mathbf{S}_B$, ρ_E is the system that the eavesdropper owns, ρ_{SE} is the classical-quantum state describing the joint state of S and E, U_S is the uniform mixture of all possible values of S, and p_{pass} is the probability that all steps of the protocol are successfully conducted.

The proof technique based on the uncertainty principle for the smooth entropy [50] can provide us an efficient method for the finite-key analysis [6]. Likely, for the protocol that we give in this paper, it can also be used as a tool for deriving the finite secret key length ℓ. For simplicity, we directly introduce the results of [44,47] into our protocol. The differences are that the key is generated from the Z basis and all of the detection events in the Z basis contribute to the final key in our protocol. Furthermore, we assume that Alice and Bob use forward classical communication [18] for error correction and privacy amplification. Then, the protocol is ε-secure, if an upper bound on ℓ can be derived as [44,47]

$$\ell \leqslant n^Z_0 + n^Z_1[1 - h(\varphi_Z)] - \lambda_{\text{EC}} - \log_2\frac{2}{\varepsilon_{\text{cor}}} - 6\log_2\frac{21}{\varepsilon_{\text{sec}}}, \tag{1}$$

where $h(x) = -x\log_2(x) - (1-x)\log_2(1-x), \varepsilon_{\text{cor}} = \varepsilon_{\text{hash}}$, $\varepsilon_{\text{cor}} + \varepsilon_{\text{sec}} \leqslant \varepsilon$, $n^Z_0 = e^{-\mu_a}n^Z_{0\mu_b} + e^{-\upsilon_a}n^Z_{0\mu_b} + e^{-\mu_a}n^Z_{0\upsilon_b} + e^{-\upsilon_a}n^Z_{0\upsilon_b}$ with $n^Z_{0\mu_b},n^Z_{0\upsilon_b}$ being the number that there is no photon from Alice's side and a successful measurement event occurs, and $n^Z_1 = \sum_{\alpha\neq 0\in\mathcal{A},\beta\neq 0\in\mathcal{B}} n^{Z,\alpha\beta}_{11}$ with $n^{Z,\alpha\beta}_{11}$ being the number of single-photon events in the Z basis when Alice and Bob send pulses of the intensity α and β, respectively. φ_Z is the phase error rate with respect to the single-photon events. $\lambda_{\text{EC}} = f\sum_{\alpha\neq 0\in\mathcal{A},\beta\neq 0\in\mathcal{B}} m^Z_{\alpha\beta}$ is the number of bits consumed in the error correction step with f being the efficiency and $m^Z_{\alpha\beta}$ being the number of bit errors in the Z basis when Alice and Bob send pulses of the intensity α and β, respectively.

IV. PARAMETER ESTIMATION IN THE FINITE-KEY CASE

In this section, our major task is to estimate the terms n^Z_1 and φ_Z in Eq. (1) using the decoy-state method. We note that the analytical security bounds derived by Yu *et al.* [30] can achieve a better secret key rate compared to existing ones. However, their result is confined to the infinite-key case. What we shall do next is to derive the bounds for the finite-key case following the approach proposed in [30].

A. Single-photon detections

The key point of the decoy-state method is that the transmitted state over the channel appears the same to the eavesdropper no matter how Alice and Bob choose the intensity level. Denote s^Z_{nm} as the successful detections in the Z basis when Alice sends n-photon state and Bob sends m-photon state. Then, we can find that $n^Z = \sum_{n,m=0}^{\infty} s^Z_{nm}$. Let $\alpha\beta$ be a two-pulse source when the pulse pair is prepared if the intensity $\alpha \in \mathcal{A}$ is chosen at Alice's side and the intensity $\beta \in \mathcal{B}$ is chosen at Bob's side. By exploiting the structure of the conditional probabilities with Baye's rule $p^Z_{\alpha\beta|nm}$ [44], we can further classify the successful detection events according to different two-pulse sources. Specifically, the number of successful detections given source $\alpha\beta$ in the asymptotic limit, which is denoted as $\tilde{n}^Z_{\alpha\beta}$, can be expressed by

$$\tilde{n}^Z_{\alpha\beta} = \sum_{n=0,m=0}^{\infty} p^Z_{\alpha\beta|nm} s^Z_{nm}, \quad \forall \alpha \in \mathcal{A}, \beta \in \mathcal{B}, \tag{2}$$

where

$$p^Z_{\alpha\beta|nm} = \frac{p_{\alpha,\beta,Z} a_n b_m}{\tau^Z_{nm}}, \quad \tau^Z_{nm} = \sum_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} p_{\alpha,\beta,Z} a_n b_m \tag{3}$$

In the above equation, $p^Z_{\alpha\beta|nm}$ is the conditional probability of choosing the source $\alpha\beta$ given that Alice and Bob prepared a n-photon state and a m-photon state, respectively, $p_{\alpha,\beta,Z}$ represents the probability that Alice chooses intensity α and Bob chooses intensity β under the Z basis, a_n denotes the probability of a n-photon pulse coming from the intensity α at Alice's side, and b_m denotes the probability of a m-photon pulse coming from the intensity β at Bob's side. In the finite-key case, if we apply Hoeffdings inequality for independent events [52], the real measured value $n^Z_{\alpha\beta}$ shall deviate from the asymptotic value $\tilde{n}^Z_{\alpha\beta}$ by

$$\left| n^Z_{\alpha\beta} - \tilde{n}^Z_{\alpha\beta} \right| \leqslant \delta(n^Z, \varepsilon_1), \tag{4}$$

with a failure probability at most $2\varepsilon_1$, where $\delta(n^Z,\varepsilon_1) = \sqrt{0.5 n^Z \ln(1/\varepsilon_1)}$ with n_Z the sum of successful detection events in the Z basis when either Alice or Bob sends no vacuum states. Essentially, Eq. (4) enables us to construct a link between the asymptotic values and the observed statistics such as $n^Z_{\mu_a\mu_b}$, $n^Z_{\mu_a\upsilon_b}$, $n^Z_{\upsilon_a\mu_b}$, and $n^Z_{\upsilon_a\upsilon_b}$.

We assume that the nonvacuum states from Alice and Bob's laboratories have the following convex forms in photon number space:

$$\begin{aligned} \rho_{\upsilon_a} &= \sum_{k=0}^{\infty} a_k \,|k\rangle\,\langle k|, \quad \rho_{\mu_a} = \sum_{k=0}^{\infty} a'_k |k\rangle\langle k|, \\ \rho_{\upsilon_b} &= \sum_{k=0}^{\infty} b_k |k\rangle\langle k|, \quad \rho_{\mu_b} = \sum_{k=0}^{\infty} b'_k |k\rangle\langle k|, \end{aligned} \tag{5}$$

with the photon number probabilities satisfying [30]

$$\frac{a'_k}{a_k} \geqslant \frac{a'_2}{a_2} \geqslant \frac{a'_1}{a_1}, \quad \frac{b'_k}{b_k} \geqslant \frac{b'_2}{b_2} \geqslant \frac{b'_1}{b_1}, \tag{6}$$

for $k \geqslant 2$. Considering the successful detections under the Z basis from source $\upsilon_a\upsilon_b$, $\upsilon_a\mu_b$, $\mu_a\upsilon_b$, and $\mu_a\mu_b$, respectively, we have the following equations:

$$\tilde{n}^{Z*}_{\upsilon_a\upsilon_b} = a_1 b_1 g_{11} + a_1 b_2 g_{12} + a_2 b_1 g_{21} + a_2 b_2 g_{22} + G_{\upsilon_a\upsilon_b}, \tag{7}$$

$$\tilde{n}^{Z*}_{\upsilon_a\mu_b} = a_1 b'_1 g_{11} + a_1 b'_2 g_{12} + a_2 b'_1 g_{21} + a_2 b'_2 g_{22} + G_{\upsilon_a\mu_b}, \tag{8}$$

$$\tilde{n}^{Z*}_{\mu_a\upsilon_b} = a'_1 b_1 g_{11} + a'_1 b_2 g_{12} + a'_2 b_1 g_{21} + a'_2 b_2 g_{22} + G_{\mu_a\upsilon_b}, \tag{9}$$

$$\tilde{n}^{Z*}_{\mu_a\mu_b} = a'_1 b'_1 g_{11} + a'_1 b'_2 g_{12} + a'_2 b'_1 g_{21} + a'_2 b'_2 g_{22} + G_{\mu_a\mu_b}, \tag{10}$$

where

$$\begin{aligned} \tilde{n}^{Z*}_{\upsilon_a\upsilon_b} &= \frac{\tilde{n}^Z_{\upsilon_a\upsilon_b}}{p_{\upsilon_a} p_{\upsilon_b}(1-q_X)} - \frac{a_0 \tilde{n}^Z_{0\upsilon_b}}{(1-p_{\upsilon_a}-p_{\mu_a}) p_{\upsilon_b}(1-q_X)} - \frac{b_0 \tilde{n}^Z_{\upsilon_a 0}}{(1-p_{\upsilon_b}-p_{\mu_b}) p_{\upsilon_a}(1-q_X)} + \frac{a_0 b_0 \tilde{n}^Z_{00}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})}, \\ \tilde{n}^{Z*}_{\upsilon_a\mu_b} &= \frac{\tilde{n}^Z_{\upsilon_a\mu_b}}{p_{\upsilon_a} p_{\mu_b}(1-q_X)} - \frac{a_0 \tilde{n}^Z_{0\mu_b}}{(1-p_{\upsilon_a}-p_{\mu_a}) p_{\mu_b}(1-q_X)} - \frac{b'_0 \tilde{n}^Z_{\upsilon_a 0}}{(1-p_{\upsilon_b}-p_{\mu_b}) p_{\upsilon_a}(1-q_X)} + \frac{a_0 b'_0 \tilde{n}^Z_{00}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})}, \\ \tilde{n}^{Z*}_{\mu_a\upsilon_b} &= \frac{\tilde{n}^Z_{\mu_a\upsilon_b}}{p_{\mu_a} p_{\upsilon_b}(1-q_X)} - \frac{a'_0 \tilde{n}^Z_{0\upsilon_b}}{(1-p_{\upsilon_a}-p_{\mu_a}) p_{\upsilon_b}(1-q_X)} - \frac{b_0 \tilde{n}^Z_{\mu_a 0}}{(1-p_{\upsilon_b}-p_{\mu_b}) p_{\mu_a}(1-q_X)} + \frac{a'_0 b_0 \tilde{n}^Z_{00}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})}, \\ \tilde{n}^{Z*}_{\mu_a\mu_b} &= \frac{\tilde{n}^Z_{\mu_a\mu_b}}{p_{\mu_a} p_{\mu_b}} - \frac{a'_0 \tilde{n}^Z_{0\mu_b}}{(1-p_{\upsilon_a}-p_{\mu_a}) p_{\mu_b}(1-q_X)} - \frac{b'_0 \tilde{n}^Z_{\mu_a 0}}{(1-p_{\upsilon_b}-p_{\mu_b}) p_{\mu_a}(1-q_X)} + \frac{a'_0 b'_0 \tilde{n}^Z_{00}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})}, \end{aligned} \tag{11}$$

and

$$G_{\upsilon_a\upsilon_b} = \sum_{(n,m)\in G_0} a_n b_m g_{nm}, \quad G_{\upsilon_a\mu_b} = \sum_{(m,n)\in G_0} a_n b'_m g_{nm}, \quad G_{\mu_a\upsilon_b} = \sum_{(n,m)\in G_0} a'_n b_m g_{nm}, \quad G_{\mu_a\mu_b} = \sum_{(n,m)\in G_0} a'_n b'_m g_{nm}, \tag{12}$$

with $g_{nm} = \frac{s^Z_{nm}}{\tau^Z_{nm}}$ $(n \geqslant 1, m \geqslant 1)$ and $G_0 = \{(n,m) | n \geqslant 1, m \geqslant 1, n+m \geqslant 4, (n,m) \neq (2,2)\}$.

One can obtain different analytical bounds of s^Z_{11} from Eqs. (7)–(10) by eliminating the terms g_{12} and g_{21}. Here, in order to obtain the most compact one among them, we apply the same analysis conducted in [30]. By comparing the coefficients of g_{nm} $(n \geqslant 2, m \geqslant 2)$ for different approaches, it can be proved that the bound derived by $(a_1 a'_2 b_1 b'_2 - a'_1 a_2 b'_1 b_2) \times$ Eq. (7) $-$ $b_1 b_2 (a_1 a'_2 - a'_1 a_2) \times$ Eq. (8) $-$ $a_1 a_2 (b_1 b'_2 - b'_1 b_2) \times$ Eq. (9) is the lowest one among them, which is presented by

$$s^Z_{11} \geqslant \frac{\tau^Z_{11}\left[(a_1 a'_2 b_1 b'_2 - a'_1 a_2 b'_1 b_2)\tilde{n}^{Z*}_{\upsilon_a\upsilon_b} - b_1 b_2 (a_1 a'_2 - a'_1 a_2)\tilde{n}^{Z*}_{\upsilon_a\mu_b} - a_1 a_2 (b_1 b'_2 - b'_1 b_2)\tilde{n}^{Z*}_{\mu_a\upsilon_b}\right]}{a_1 b_1 (a_1 a'_2 - a'_1 a_2)(b_1 b'_2 - b'_1 b_2)}, \tag{13}$$

where $\tilde{n}^{Z*}_{\upsilon_a\upsilon_b}$, $\tilde{n}^{Z*}_{\upsilon_a\mu_b}$, and $\tilde{n}^{Z*}_{\mu_a\upsilon_b}$ are determined by Eq. (11). Particularly, it should be noted that Eq. (6) is a necessary condition for the correctness of the above bound. Furthermore, consider the finite-size deviation terms given by Eq. (4); we can reformulate the bound with experimental observed events by

$$s_{11}^{Z} \geqslant \frac{\tau_{11}^{Z}\left[(a_1a_2'b_1b_2' - a_1'a_2b_1'b_2)N_{\upsilon_a\upsilon_b}^{Z} - b_1b_2(a_1a_2' - a_1'a_2)N_{\upsilon_a\mu_b}^{Z} - a_1a_2(b_1b_2' - b_1'b_2)N_{\mu_a\upsilon_b}^{Z}\right]}{a_1b_1(a_1a_2' - a_1'a_2)(b_1b_2' - b_1'b_2)} \triangleq \underline{s_{11}^{Z}}, \tag{14}$$

where

$$\begin{aligned} N_{\upsilon_a\upsilon_b}^{Z} &= \frac{n_{\upsilon_a\upsilon_b}^{Z} - \delta(n^Z,\varepsilon_1)}{p_{\upsilon_a}p_{\upsilon_b}(1-q_X)} - \frac{a_0\left[n_{0\upsilon_b}^{Z} + \delta(n^Z,\varepsilon_1)\right]}{(1-p_{\upsilon_a}-p_{\mu_a})p_{\upsilon_b}(1-q_X)} - \frac{b_0\left[n_{\upsilon_a 0}^{Z} + \delta(n^Z,\varepsilon_1)\right]}{(1-p_{\upsilon_b}-p_{\mu_b})p_{\upsilon_a}(1-q_X)} + \frac{a_0b_0n_{00}^{Z}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})},\\ N_{\upsilon_a\mu_b}^{Z} &= \frac{n_{\upsilon_a\mu_b}^{Z} + \delta(n^Z,\varepsilon_1)}{p_{\upsilon_a}p_{\mu_b}(1-q_X)} - \frac{a_0\left[n_{0\mu_b}^{Z} - \delta(n^Z,\varepsilon_1)\right]}{(1-p_{\upsilon_a}-p_{\mu_a})p_{\mu_b}(1-q_X)} - \frac{b_0'\left[n_{\upsilon_a 0}^{Z} - \delta(n^Z,\varepsilon_1)\right]}{(1-p_{\upsilon_b}-p_{\mu_b})p_{\upsilon_a}(1-q_X)} + \frac{a_0b_0'n_{00}^{Z}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})},\\ N_{\mu_a\upsilon_b}^{Z} &= \frac{n_{\mu_a\upsilon_b}^{Z} + \delta(n^Z,\varepsilon_1)}{p_{\mu_a}p_{\upsilon_b}(1-q_X)} - \frac{a_0'\left[n_{0\upsilon_b}^{Z} - \delta(n^Z,\varepsilon_1)\right]}{(1-p_{\upsilon_a}-p_{\mu_a})p_{\upsilon_b}(1-q_X)} - \frac{b_0\left[n_{\mu_a 0}^{Z} - \delta(n^Z,\varepsilon_1)\right]}{(1-p_{\upsilon_b}-p_{\mu_b})p_{\mu_a}(1-q_X)} + \frac{a_0'b_0n_{00}^{Z}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})}, \end{aligned} \tag{15}$$

with $n_{\alpha\beta}^{Z}$ being the number of measured successful detections from source $\alpha\beta$ in the Z basis.

B. Single-photon errors

In what follows, we shall first derive a bound on how to calculate the single-photon errors based on the observed errors in the X basis. In our protocol, to obtain a higher sifted key, we apply a biased basis choice. Most importantly, for the states prepared in the X basis from Alice and Bob's laboratories, only the decoy source $\upsilon_a\upsilon_b$ is selected for parameter estimation. Denote the υ_{nm}^{X} as the number of errors associated with s_{nm}^{X}, which is the number of successful detections in the X basis when Alice sends n-photon state and Bob sends m-photon state. Then, the expected number of errors assigned to the source $\upsilon_a\upsilon_b$ can be represented by

$$\tilde{w}_{\upsilon_a\upsilon_b}^{X} = \sum_{n=0,m=0}^{\infty} p_{\alpha\beta|nm}^{X}\upsilon_{nm}^{X}, \tag{16}$$

where

$$p_{\alpha\beta|nm}^{X} = \frac{p_{\alpha,\beta,X}a_nb_m}{\tau_{nm}^{X}}, \quad \tau_{nm}^{X} = \sum_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} p_{\alpha,\beta,X}a_nb_m. \tag{17}$$

Also using Hoeffdings inequality [52], we thus obtain the real measured errors $w_{\upsilon_a\upsilon_b}^{X}$ which deviates from the expected ones by

$$\left|w_{\upsilon_a\upsilon_b}^{X} - \tilde{w}_{\upsilon_a\upsilon_b}^{X}\right| \leqslant \delta(w^X,\varepsilon_2), \tag{18}$$

which holds true with a probability of at least $1-2\varepsilon_2$. Here, $\delta(w^X,\varepsilon_w) = \sqrt{0.5w^X\ln(1/\varepsilon_2)}$, where with w^X is the sum of errors in the X basis when either Alice or Bob sends no vacuum states.

According to the conditional probabilities defined in the previous section, we can obtain the expression of asymptotical errors from source $\upsilon_a\upsilon_b$ in photon-number distribution

$$\tilde{w}_{\upsilon_a\upsilon_b}^{X*} = a_1b_1r_{11} + a_1b_2r_{12} + a_2b_1r_{21} + a_2b_2r_{22} + T_{\upsilon_a\upsilon_b}, \tag{19}$$

where

$$\begin{aligned} \tilde{w}_{\upsilon_a\upsilon_b}^{X*} &= \frac{\tilde{w}_{\upsilon_a\upsilon_b}^{X}}{p_{\upsilon_a}p_{\upsilon_b}q_X} - \frac{a_0\tilde{w}_{0\upsilon_b}^{X}}{(1-p_{\upsilon_a}-p_{\mu_a})p_{\upsilon_b}q_X}\\ &\quad - \frac{b_0\tilde{w}_{\upsilon_a 0}^{X}}{(1-p_{\upsilon_b}-p_{\mu_b})p_{\upsilon_a}q_X}\\ &\quad + \frac{a_0b_0\tilde{w}_{00}^{X}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})},\\ T_{\upsilon_a\upsilon_b} &= \sum_{(n,m)\in G_0} a_nb_mr_{nm}, \end{aligned} \tag{20}$$

with $r_{nm} = \frac{\upsilon_{nm}^{X}}{\tau_{nm}^{X}}$ $(n \geqslant 1, m \geqslant 1)$.

Note that all terms in the right-hand side of Eq. (19) are not smaller than zero and ignore the terms when either of the subscripts n or m is bigger than 1; we can obtain an upper bound of υ_{11}^{X}, which is given by

$$\upsilon_{11}^{X} \leqslant \frac{\tau_{11}^{X}\tilde{w}_{\upsilon_a\upsilon_b}^{X*}}{a_1b_1}. \tag{21}$$

Then, consider the deviation terms between the number of expected errors and observed errors; we can obtain

$$\upsilon_{11}^{X} \leqslant \frac{\tau_{11}^{X}W_{\upsilon_a\upsilon_b}^{X}}{a_1b_1} \triangleq \overline{\upsilon_{11}^{X}}, \tag{22}$$

with

$$\begin{aligned} W_{\upsilon_a\upsilon_b}^{X} &= \frac{w_{\upsilon_a\upsilon_b}^{X} + \delta(w^X,\varepsilon_2)}{p_{\upsilon_a}p_{\upsilon_b}q_X} - \frac{a_0\left[w_{0\upsilon_b}^{X} - \delta(w^X,\varepsilon_2)\right]}{(1-p_{\upsilon_a}-p_{\mu_a})p_{\upsilon_b}q_X}\\ &\quad - \frac{b_0\left[w_{\upsilon_a 0}^{X} - \delta(w^X,\varepsilon_2)\right]}{(1-p_{\upsilon_b}-p_{\mu_b})p_{\upsilon_a}q_X}\\ &\quad + \frac{a_0b_0w_{00}^{X}}{(1-p_{\upsilon_a}-p_{\mu_a})(1-p_{\upsilon_b}-p_{\mu_b})}. \end{aligned} \tag{23}$$

C. Phase error rate

Since the final secret key is produced in the Z basis, we need to the estimate the phase error rate under the Z basis, i.e., φ_Z. In the asymptotic case, the phase error rate in the Z basis is equal to the bit error rate in the X basis. However, in the finite-key case, φ_Z should be estimated via a random-sampling theory (without replacement) based on the observed error rate in the X basis. Since the security of MDI-QKD is built on the time-reversed Einstein-Podolsky-Rosen protocols [53,54], the analysis on estimating the phase error rate for BB84 protocol can be naturally introduced into that for MDI-QKD. In this paper, we conduct an interval estimation based on the straightforward bounds [36], derived from an approaching technique for the hypergeometric distribution. And we stress that the estimation method given in [55] is also preferable in our protocol. It should be noted that the result of [36] is obtained for the ideal single-photon source. Here, combining the decoy-state method with the bound given by Hayashi [36], we will show how to calculate φ_Z for the case when nonideal single-photon sources are used, such as the WCS and the heralded SPDCS.

First of all, we should give an estimation on the number of sample single-photon detections in the X basis. The method is similar to the way of calculating $\underline{s}_{11}^{Z}$ and the calculation formula can be given by

$$s_{11}^{X} \geqslant \frac{\tau_{11}^{X}\left[(a_1a_2'b_1b_2' - a_1'a_2b_1'b_2)N_{\upsilon_a\upsilon_b}^{X} - b_1b_2(a_1a_2' - a_1'a_2)N_{\upsilon_a\mu_b}^{X} - a_1a_2(b_1b_2' - b_1'b_2)N_{\mu_a\upsilon_b}^{X}\right]}{a_1b_1(a_1a_2' - a_1'a_2)(b_1b_2' - b_1'b_2)} \triangleq \underline{s}_{11}^{X}, \tag{24}$$

where

$$\begin{aligned}
N_{\upsilon_a\upsilon_b}^{X} &= \frac{n_{\upsilon_a\upsilon_b}^{X} - \delta(n^X,\varepsilon_1)}{p_{\upsilon_a}p_{\upsilon_b}q_X} - \frac{a_0\left[n_{0\upsilon_b}^{X} + \delta(n^X,\varepsilon_1)\right]}{\left(1 - p_{\upsilon_a} - p_{\mu_a}\right)p_{\upsilon_b}q_X} - \frac{b_0\left[n_{\upsilon_a0}^{X} + \delta(n^X,\varepsilon_1)\right]}{\left(1 - p_{\upsilon_b} - p_{\mu_b}\right)p_{\upsilon_a}q_X} + \frac{a_0b_0n_{00}^{X}}{\left(1 - p_{\upsilon_a} - p_{\mu_a}\right)\left(1 - p_{\upsilon_b} - p_{\mu_b}\right)},\\
N_{\upsilon_a\mu_b}^{X} &= \frac{n_{\upsilon_a\mu_b}^{X} + \delta(n^X,\varepsilon_1)}{p_{\upsilon_a}p_{\mu_b}q_X} - \frac{a_0\left[n_{0\mu_b}^{X} - \delta(n^X,\varepsilon_1)\right]}{\left(1 - p_{\upsilon_a} - p_{\mu_a}\right)p_{\mu_b}q_X} - \frac{b_0'\left[n_{\upsilon_a0}^{X} - \delta(n^X,\varepsilon_1)\right]}{\left(1 - p_{\upsilon_b} - p_{\mu_b}\right)p_{\upsilon_a}q_X} + \frac{a_0b_0'n_{00}^{X}}{\left(1 - p_{\upsilon_a} - p_{\mu_a}\right)\left(1 - p_{\upsilon_b} - p_{\mu_b}\right)},\\
N_{\mu_a\upsilon_b}^{X} &= \frac{n_{\mu_a\upsilon_b}^{X} + \delta(n^X,\varepsilon_1)}{p_{\mu_a}p_{\upsilon_b}q_X} - \frac{a_0'\left[n_{0\upsilon_b}^{X} - \delta(n^X,\varepsilon_1)\right]}{\left(1 - p_{\upsilon_a} - p_{\mu_a}\right)p_{\upsilon_b}q_X} - \frac{b_0[n_{\mu_a0}^{X} - \delta(n^X,\varepsilon_1)]}{\left(1 - p_{\upsilon_b} - p_{\mu_b}\right)p_{\mu_a}q_X} + \frac{a_0'b_0n_{00}^{X}}{\left(1 - p_{\upsilon_a} - p_{\mu_a}\right)\left(1 - p_{\upsilon_b} - p_{\mu_b}\right)},
\end{aligned} \tag{25}$$

with $n_{\alpha\beta}^{X}$ being the number of measured successful detections from source $\alpha\beta$ in the X basis and $n_{00}^{X} = n_{00}q_X$. Here, n_X is the sum of successful detection events in the X basis when either Alice or Bob send no vacuum states.

In our protocol, we note that the estimated lengths of single-photon sifted bits and single-photon sample bits are $\underline{s}_{11}^{Z}$ and $\underline{s}_{11}^{X}$, respectively. In the following, we will show how to calculate the phase error rate given $\underline{s}_{11}^{Z}$, $\underline{s}_{11}^{X}$, and $\overline{\upsilon}_{11}^{X}$.

Step 1. For a given security parameter ε_{sec}, choose a value ω such that

$$\sqrt{\frac{\underline{s}_{11}^{Z} + \underline{s}_{11}^{X}}{\underline{s}_{11}^{Z}}}\sqrt{\frac{\omega^2 + 2\pi}{2}}e^{\nu}\Phi(\omega) \leqslant \frac{1}{16}{\varepsilon_{\text{sec}}}^2, \tag{26}$$

where $\nu = \frac{1}{6\underline{s}_{11}^{Z}} + \frac{1}{12}$ and $\Phi(\omega) = \frac{1}{\sqrt{2\pi}}\int_{\omega}^{\infty}\exp(\frac{-y^2}{2})dy$.

Step 2. Determine the sample single-photon bit error rate by $e_{11}^{X} = (\overline{\upsilon}_{11}^{X} + 2)/\underline{s}_{11}^{X}$ and calculate the parameters

$$\begin{aligned}
\sigma &= \frac{\omega^2\underline{s}_{11}^{Z}}{4\underline{s}_{11}^{X}\left(\underline{s}_{11}^{Z} + \underline{s}_{11}^{X} - 1\right)},\\
\hat{e} &= \frac{e_{11}^{X} + 2\sigma + 2\sqrt{\sigma\left[e_{11}^{X}\left(1 - e_{11}^{X}\right) + \sigma\right]}}{1 + 4\sigma}.
\end{aligned} \tag{27}$$

Step 3. Calculate the phase error rate e_p by

$$e_p = \frac{\left(\underline{s}_{11}^{Z} + \underline{s}_{11}^{X}\right)\hat{e} - \underline{s}_{11}^{X}e_{11}^{X}}{\underline{s}_{11}^{Z}}. \tag{28}$$

V. NUMERICAL SIMULATION

In this section, we will numerically simulate the performance of our protocol. For better comparison, we consider a fiber-based channel model and directly borrow experimental parameters from [44]. Denote $\eta_a^c = 10^{-\alpha L_a/10}$ ($\eta_b^c = 10^{-\alpha L_b/10}$) as the fiber transmission of Alice (Bob) with the loss coefficient being $\alpha = 0.2$ dB/km, $\eta_d = 14.5\%$ as the detection efficiency of the relay, and $p_d = 6.02 \times 10^{-6}$ as its background dark count rate. L_a (L_b) is the length of fiber between Charlie and Alice (Bob). The security bound is fixed to $\varepsilon = 10^{-10}$ and the numerical parameters are listed in Table II.

It should be noted that our protocol is not confined with specific source and any source satisfying Eq. (6) is available. Here, we consider the cases when WCS is used as the sources of Alice and Bob in the phase-encoded MDI-QKD scheme proposed by Ma *et al.* [18]. And we stress that the original polarization-based protocol [13] and the phase-encoded protocol proposed by Tamaki *et al.* [17] are also acceptable for our protocol.

TABLE II. List of experimental parameters used in the simulations: α is the loss coefficient of the fiber, f is the error correction inefficiency, η_d is the detection efficiency of the relay, e_d is the errors due to channel relative-phase misalignment between Alice and Bob, p_d is the background dark count rate of the detector in the relay, and ε is the predetermined security bound.

α (dB/km)	f	η_d	e_d	p_d	ε
0.20	1.16	0.145	0.015	6.02×10^{-6}	10^{-10}

For parameter estimation, the parameters $n^Z_{\alpha\beta}$, $n^X_{\alpha\beta}$, and $w^X_{\alpha\beta}$ can be directly obtained in the experimental process of executing our protocol. In this paper, for simulation purposes, we can theoretically calculate them based on the channel model [18]:

$$\begin{aligned} n^X_{\alpha\beta} &= 2Np_\alpha p_\beta q_X y^2[1+2y^2-4yI_0(x)+I_0(2x)],\\ w^X_{\alpha\beta} &= e_0 n^X_{\alpha\beta} - 2Np_\alpha p_\beta q_X(e_0-e_d)y^2[I_0(2x)-1],\\ n^Z_{\alpha\beta} &= Np_\alpha p_\beta(1-q_X)(Q_C+Q_E),\\ w^Z_{\alpha\beta} &= Np_\alpha p_\beta(1-q_X)\{e_dQ_C+(1-e_d)Q_E\}, \end{aligned} \tag{29}$$

where

$$\begin{aligned} Q_C &= 2(1-p_d)^2 e^{-\mu'/2}[1-(1-p_d)e^{-\eta_a\alpha/2}]\\ &\quad\times[1-(1-p_d)e^{-\eta_b\beta/2}], \\ Q_E &= 2p_d(1-p_d)^2e^{-\mu'/2}[I_0(2x)-(1-p_d)e^{-\mu'/2}], \end{aligned} \tag{30}$$

with $\mu' = \eta_a\alpha + \eta_b\beta$, $x = \sqrt{\eta_a\alpha\eta_b\beta}/2$, and $y = (1-p_d)e^{-\mu'/4}$. In the above equations, $I_0(x)$ is the modified Bessel function of the first kind, $e_0 = 0.5$ is the error rate of a random background noise, p_d is the dark count rate of the detector in the relay, and $\eta_a = \eta_d\eta^c_a$ $(\eta_b = \eta_d\eta^c_b)$ is the transmission efficiency of Alice (Bob).

In our simulations, we use the experimental parameters listed in Table II. We set the key's secrecy ε_{sec} and correctness ε_{cor} be 10^{-10} and 10^{-12}, respectively.

In Fig. 1, the curves of the numerical secret key rates from left to right are obtained for different values of q_X with the total number of signals N fixed to be 10^{13}. One can see that the secret key rate and achievable secure transmission distance have their optimal values for different basis choices. 0.25 is found to be optimal for $N = 10^{13}$. This comes from the trade-off between q_X and q_Z since the bigger q_X means less statistical fluctuation and the bigger $q_Z = 1 - q_X$ means more generation of key.

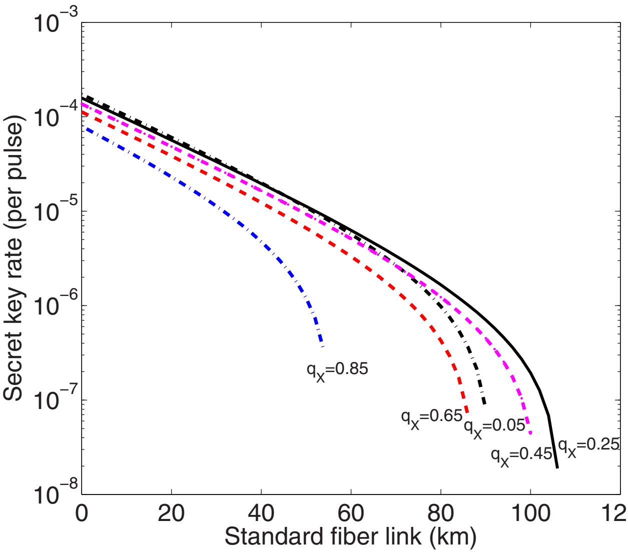

FIG. 1. (Color online) Key rate comparison with different probabilities of choosing X basis. A realistic finite size of data N is fixed to be 10^{13}. For different $q_X = 0.05, 0.25, 0.45, 0.65, 0.85$, the curves from left to right are all plotted by Eq. (1) averaged on the total number of signals N sent by Alice and Bob. The single-photon contribution n^Z_1, single-photon errors υ^X_{11}, and the phase error rate φ_Z are all analytically obtained by Eq. (14), Eq. (22), and Eq. (28), respectively. The intensities of signal states and decoy states are assigned specific values of 0.36 and 0.18, respectively.

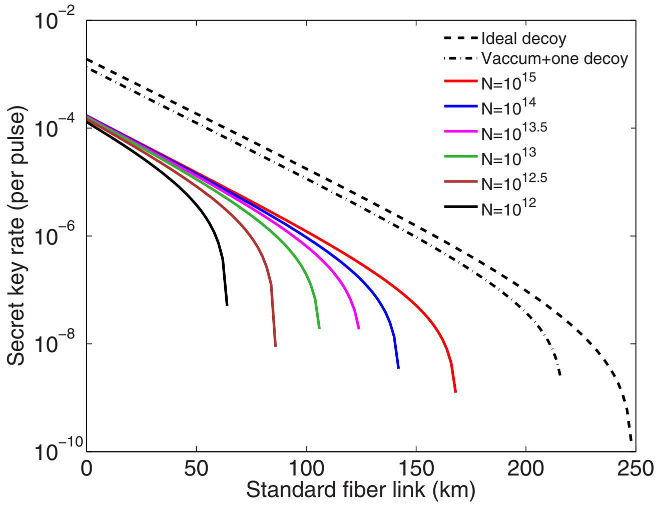

FIG. 2. (Color online) Key rate comparison with different finite data sets. The dashed curve denotes the ideal secret key rate when infinite decoy states are used and the intensity is optimized for each distance. The dash-dotted curve denotes the asymptotic secret key rate of the three-intensity decoy-state method with the intensity of signal states and decoy states being 0.5 and 0.1, respectively. The solid curves from left to right are plotted by our protocol for finite number of total pulses from the weak coherent source $N = 10^j$ with $j = 12, 12.5, 13, 13.5, 14, 15$. The intensities of signal states and decoy states are fixed to be 0.36 and 0.18, respectively. But the probability of choosing X basis q_X is optimally chosen for each N.

In Fig. 2, the solid curves of the secret key rates from left to right are obtained for different total number of signals N, where q_X is optimally chosen for each N. Under the same experimental parameters, our simulation results perform better than that of [44]. For example, with a realistic finite size of data saying $N = 10^{12}$ and $N = 10^{12.5}$, the maximal transmission distance can reach to 64 and 86 km, apparently bigger than [44]. And it should be noted that the key rates of [44] are obtained under a full optimization over various parameters such as μ_a, υ_a, p_{μ_a}, p_{υ_a}, and q_X, where $\mu_b = \mu_a$, $\upsilon_b = \upsilon_a$, $p_{\mu_b} = p_{\mu_a}$, $p_{\upsilon_b} = p_{\upsilon_a}$, and $q_Z = 1 - q_X$. However, in the simulations of our protocol, we fix $\mu_a = 0.36, \upsilon_a = 0.18, p_{\mu_a} = 0.58, p_{\upsilon_a} = 0.3$ and only optimize the key rate over one parameter q_X. This indicates the feasibility of our protocol.

With the full parameter optimization method based on a local search algorithm [45], the key rate and secure distance of our protocol can be further increased and extended, respectively. But in practice, the implementation of full parameter optimization needs complex modulation setups, which raises the technical difficulty in practical experiment. Our protocol can be easily conducted in existing setups of a practical experiment. Besides, the estimation of phased error rate is conducted rigorously, whereas [44] and [45] are not. This will certainly enhance the security level of our protocol. Our protocol has good performance for two reasons. One is that we apply a smart intensity-choosing policy, where all intensity levels in the Z basis contribute to the final keys and the signal states are not prepared in the X basis when both Alice and Bob choose signal states. The other is that the basis is chosen optimally, which brings a good balance between data used for key generation and that used for parameter estimation.

VI. CONCLUSION

In conclusion, we put forward a biased decoy-state protocol for the measurement-device-independent quantum key distribution with finite resources. For secret keys of a finite length that are ε-secure, the analytical bounds for the single-photon detections, single-photon errors, and phase error rate are presented. Our protocol takes advantage of all possible detections even if Alice and Bob choose decoy states. The basis choice is assigned according to the intensity choice. By numerical simulations, we remark that biased basis choice can apparently extend the secure distance for MDI-QKD. Without a full optimization program, our protocol can perform better than that proposed by Curty *et al.* [44] and the secure distance can reach to 168 km when the size of total pulses is $N = 10^{15}$. Experimentalists can readily choose and implement our protocol to enhance the performance of practical MDI-QKD experiment.

Recently, we noticed that a study by Zhou *et al.* [46] proposed analytical estimation bounds based on four intensities for MDI-QKD. It will be appealing to adjust these bounds to our protocol and we will maintain attention to the new application of our protocol.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support from the National Basic Research Program of China (Grant No. 2013CB338002) and the National Natural Science Foundation of China (Grant No. 11304397).

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

[2] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[3] J. Cederlöf and J.-Ä. Larsson, IEEE Trans. Inf. Theory **54**, 1735 (2008).

[4] C. Zhou, W.-S. Bao, H.-W. Li, Y. Wang, and X.-Q. Fu, Quantum. Inf. Process. **13**, 935 (2014).

[5] V. Scarani and R. Renner, in *Theory of Quantum Computation, Communication, and Cryptography*, edited by Y. Kawano and M. Mosca, Lecture Notes in Computer Science Vol. 5106 (Springer, Berlin/Heidelberg, 2008), pp. 83–95; V. Scarani and R. Renner, Phys. Rev. Lett. **100**, 200501 (2008).

[6] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, Nat. Commun. **3**, 634 (2012).

[7] Y. Zhao, Chi-Hang Fred Fung, B. Qi, C. Chen, and H.-K. Lo, Phys. Rev. A **78**, 042333 (2008).

[8] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Nat. Photonics **4**, 686 (2010).

[9] H.-W. Li *et al.*, Phys. Rev. A **84**, 062308 (2011).

[10] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).

[11] N. Gisin, S. Pironio, and N. Sangouard, Phys. Rev. Lett. **105**, 070501 (2010).

[12] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[13] H.-K. Lo, M. Curty, and B. Qi, Phys. Rev. Lett. **108**, 130503 (2012).

[14] S. L. Braunstein and S. Pirandola, Phys. Rev. Lett. **108**, 130502 (2012).

[15] M. Pawłowski and N. Brunner, Phys. Rev. A **84**, 010302 (2011).

[16] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, Phys. Rev. A **85**, 010301 (2012).

[17] K. Tamaki, H.-K. Lo, Chi-Hang Fred Fung, and B. Qi, Phys. Rev. A **85**, 042307 (2012).

[18] X. F. Ma and M. Razavi, Phys. Rev. A **86**, 062319 (2012).

[19] X.-B. Wang, Phys. Rev. A **87**, 012320 (2013).

[20] Z.-Q. Yin, Chi-Hang Fred Fung, X. Ma, C.-M. Zhang, H.-W. Li, W. Chen, S. Wang, G.-C. Guo, and Z.-F. Han, Phys. Rev. A **88**, 062322 (2013).

[21] Y. Liu, T.-Y. Chen, L.-J. Wang, H. Liang, G.-L. Shentu, J. Wang, Ke Cui, Hua-Lei Yin, Nai-Le Liu, Li Li *et al.*, Phys. Rev. Lett. **111**, 130502 (2013).

[22] A. Rubenok, J. A. Slater, P. Chan, I. Lucio-Martinez, and W. Tittel, Phys. Rev. Lett. **111**, 130501 (2013).

[23] Z. Tang, Z. Liao, F. Xu, B. Qi, L. Qian, and H.-K. Lo, Phys. Rev. Lett. **112**, 190503 (2014).

[24] W.-Y. Hwang, Phys. Rev. Lett. **91**, 057901 (2003).

[25] H.-K. Lo, X. Ma, and K. Chen, Phys. Rev. Lett. **94**, 230504 (2005).

[26] X.-B. Wang, Phys. Rev. Lett. **94**, 230503 (2005).

[27] S.-H. Sun, M. Gao, C.-Y. Li, and L.-M. Liang, Phys. Rev. A **87**, 052329 (2013).

[28] Q. Wang and X.-B. Wang, Phys. Rev. A **88**, 052332 (2013).

[29] C. Zhou, W.-S. Bao, W. Chen, H.-W. Li, Z.-Q. Yin, Y. Wang, and Z.-F. Han, Phys. Rev. A **88**, 052333 (2013).

[30] Z.-W. Yu, Y.-H. Zhou, and X.-B. Wang, Phys. Rev. A **88**, 062339 (2013).

[31] M. Hayashi, Phys. Rev. A **76**, 012329 (2007); J. Hasegawa, M. Hayashi, T. Hiroshima, and A. Tomita, arXiv:0707.3541.

[32] H.-W. Li, Y.-B. Zhao, Z.-Q. Yin, S. Wang, Z.-F. Han, W.-S. Bao, and G.-C. Guo, Opt. Commun. **282**, 4162 (2009).

[33] M. Christandl, R. König, and R. Renner, Phys. Rev. Lett. **102**, 020504 (2009).

[34] T.-T. Song, J. Zhang, S.-J. Qin, and Q.-Y. Wen, Quantum Inf. Comput. **11**, 374 (2011).

[35] Nelly Huei Ying Ng, M. Berta, and S. Wehner, Phys. Rev. A **86**, 042315 (2012).

[36] M. Hayashi and T. Tsurumaru, New J. Phys. **14**, 093014 (2012).

[37] M. Hayashi and R. Nakayama, New J. Phys. **16**, 063009 (2014).

[38] R. D. Somma and R. J. Hughes, Phys. Rev. A **87**, 062330 (2013).

[39] Y. Wang, W.-S. Bao, H.-W. Li, C. Zhou, and Y. Li, Phys. Rev. A **88**, 052322 (2013).

[40] C. Zhou, W.-S. Bao, H.-W. Li, Y. Wang, Y. Li, Z.-Q. Yin, W. Chen, and Z.-F. Han, Phys. Rev. A **89**, 052328 (2014).

[41] M. Mafu, K. Garapo, and F. Petruccione, Phys. Rev. A **88**, 062306 (2013).

[42] T.-T. Song, Q.-Y. Wen, F.-Z. Guo, and X.-Q. Tan, Phys. Rev. A **86**, 022332 (2012).

[43] X. Ma, Chi-Hang Fred Fung, and M. Razavi, Phys. Rev. A **86**, 052305 (2012).

[44] M. Curty, F.-H. Xu, W. Cui, C. C. W. Lim, K. Tamaki, and H.-K. Lo, Nat. Commun. **5**, 3732 (2014).

[45] F.-H. Xu, H. Xu, and H.-K. Lo, Phys. Rev. A **89**, 052333 (2014).

[46] Y.-H. Zhou, Z.-W. Yu, and X.-B. Wang, Phys. Rev. A **89**, 052325 (2014).

[47] Charles Ci Wen Lim, M. Curty, N. Walenta, F. Xu, and H. Zbinden, Phys. Rev. A **89**, 022307 (2014).

[48] M. Lucamarini, K. A. Patel, J. F. Dynes, B. Fröhlich, A. W. Sharpe, A. R. Dixon, Z. L. Yuan, R. V. Penty, and A. J. Shields, Opt. Express **21**, 24550 (2013).

[49] Z. Wei, W. Wang, Z. Zhang, M. Gao, Z. Ma, and X. Ma, Sci. Rep. **3**, 2453 (2013).

[50] M. Tomamichel and R. Renner, Phys. Rev. Lett. **106**, 110506 (2011).

[51] R. Renner, Ph.D. thesis, Swiss Federal Institute of Technology (ETH) Zurich, 2005; Int. J. Quantum Inf. **6**, 1 (2008).

[52] W. Hoeffding, J. Am. Stat. Assoc. **58**, 13 (1963).

[53] E. Biham, B. Huttner, and T. Mor, Phys. Rev. A **54**, 2651 (1996).

[54] H. Inamori, Algorithmica **34**, 340 (2002).

[55] Chi-Hang Fred Fung, X. Ma, and H. F. Chau, Phys. Rev. A **81**, 012318 (2010).