Loss-tolerant quantum cryptography with imperfect sources

Kiyoshi Tamaki,[1,*] Marcos Curty,[2] Go Kato,[3] Hoi-Kwong Lo,[4] and Koji Azuma[1]

[1]*NTT Basic Research Laboratories, NTT Corporation, 3-1, Morinosato Wakamiya Atsugi-Shi, Kanagawa, 243-0198, Japan*

[2]*EI Telecomunicación, Department of Signal Theory and Communications, University of Vigo, E-36310, Spain*

[3]*NTT Communication Science Laboratories, NTT Corporation, 3-1, Morinosato Wakamiya Atsugi-Shi, Kanagawa, 243-0198, Japan*

[4]*Center for Quantum Information and Quantum Control, Department of Electrical & Computer Engineering and Department of Physics, University of Toronto, Canada M5S 3G4*

(Received 26 December 2013; revised manuscript received 17 September 2014; published 12 November 2014)

In principle, quantum key distribution (QKD) offers unconditional security based on the laws of physics. Unfortunately, all previous QKD experiments assume perfect state preparation in their security analysis. Therefore, the generated key is *not* proven to be secure in the presence of unavoidable modulation errors. The key reason that modulation errors are not considered in previous QKD experiments lies in a crucial weakness of the standard Gottesman-Lo-Lütkenhaus-Preskill (GLLP) model, namely, it is not loss tolerant and Eve may in principle enhance imperfections through losses. Here, we propose a QKD protocol that is loss tolerant to state preparation flaws. Importantly, we show conclusively that the state preparation process in QKD can be much less precise than initially thought. Our method can also be applied to other quantum cryptographic protocols.

DOI: 10.1103/PhysRevA.90.052314 PACS number(s): 03.67.Dd, 03.67.Hk

I. INTRODUCTION

Quantum key distribution (QKD) [1] allows two distant parties, Alice and Bob, to share a secret key. This field has progressed very rapidly over the last years, and it now offers practical systems that can operate in realistic environments [2,3]. In principle, QKD is unconditionally secure [4–11], i.e., its security does not depend on the computational power and technologies of the eavesdropper, Eve. In practice, however, there is a gap (or so-called security loopholes) between the security proofs of QKD and its practical implementations. This is so because the photon-detection unit and the sending device do not necessarily operate as the security proofs require. Fortunately, thanks to the recent proposal of measurement-device-independent QKD (MDIQKD) [12–14], *any* security loophole in the photon-detection unit is now closed. If MDIQKD is widely deployed, which seems to be likely [15], it will shift the focus of quantum hackers to attacking the source. Here we address this important problem of securing the source in QKD.

Most QKD protocols, such as the Bennett-Brassard 1984 (BB84) [16] and the six-state [17] protocols (in either prepare-and-measure or in their MDI version), encode the key information in the state of single photons. Ideally, therefore, the sending device has to satisfy two conditions: it should emit only single photons, and it should encode the information without introducing errors. Unfortunately, however, a practical sending device, such as a phase-randomized weak coherent pulse source, fails to satisfy those requirements. That is, it can emit multiphoton pulses and, due to inevitable imperfections of phase, amplitude, and polarization modulators, the state preparation process is flawed, e.g., the polarization or the relative phase of the pulses is different from the ideal one. When the state of the single-photon part is independent of the amplitude of the coherent pulse, the first imperfection, i.e., multiphoton emissions of the source, can be solved by combining the security analysis of Gottesman, Lo, Lütkenhaus, and Preskill (GLLP) [18] with the decoy-state method [19] (henceforth referred to as the GLLP decoy). It provides a tight estimation of the detection rate of the single-photon part (i.e., the single-photon pulses emitted by Alice), together with its bit error rate. As a result, both the achievable rate and distance are significantly improved. GLLP propose as well a technique to deal with the second type of source imperfection, namely, state preparation flaws (SPFs) in the single-photon emissions. This method (henceforth referred to as the GLLP-SPF method) can be combined as well with the GLLP decoy. It has, however, a severe limitation: it is not loss tolerant. That is, the GLLP method pessimistically assumes that Eve can enhance the flaw in the single-photon part by exploiting the losses of that part, which leads to a severe degradation of the system performance; e.g., see Fig. 1. Unfortunately, one can see this degradation with the methods proposed in [10,22] as well. Thus, the claimed security of the key generated in all existing QKD experiments is not fully justified as they do not consider this flaw, which is a crucial problem in the field.

In this paper, we provide a general technique that applies to *any* SPF in phase-randomized sources. In particular, when our technique is applied to QKD implementations with decoy states, it makes the key from such QKD implementations secure again. In sharp contrast to the GLLP-SPF method, our method is loss tolerant; e.g., see Fig. 2 where we consider modulation errors just as an example of SPFs. Our technique is compatible with the use of the GLLP decoy, and it can be employed for both MDIQKD and prepare-and-measure schemes. When it is combined with MDIQKD it provides a QKD system that takes both modulation and detector flaws into account, yet the secret key rate remains almost the same [23]. Therefore, we hope that our analysis will become a standard tool for future QKD experiments. Our work builds on a simple observation: the phase error rate, a quantity that measures the information leakage to Eve in Koashi's proof [11], can be obtained by estimating the transmission

[*] tamaki.kiyoshi@lab.ntt.co.jp

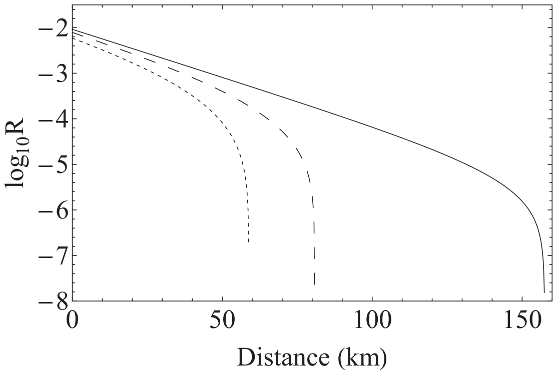

FIG. 1. Lower bound on the secret key rate R of the decoy-state BB84 protocol based on the GLLP-SPF method for different values of δ, which is the phase modulation deviation from the intended value. This is an example of SPFs, and we consider experimental parameters from [20]. The solid, dashed, and dotted lines correspond, respectively, to the cases $\delta = 0$, 0.063 (which is equivalent to 3.62°), and 0.126. The case $\delta \geqslant 0.063$ corresponds to an experimentally available value [21]. For each line, we optimize the intensity of the signals to maximize the key rate. One can see a significant decrease of the key rates when the modulation errors increase.

rate of a fictitious quantum signal sent by Alice. To obtain the rate, we exploit the basis mismatch events [24] ("rejected-data analysis" in [24] was not loss tolerant but we make it loss tolerant).

In so doing, we can (i) dramatically improve the key rate and achievable distance of QKD with SPFs; (ii) show that the three-state scheme [25,26] gives precisely the same key rate as the BB84 protocol [16]. This result is surprising, as it implies that one of the signals sent in BB84 [16] is actually redundant [27]. In addition, our technique is (iii) applicable to MDIQKD; (iv) applicable to many QKD schemes with phase-randomized sources (QKD-PRS schemes) (it can be shown, for instance, that a particular four-state scheme can postprocess its data following the specifications of the six-state protocol [17] to outperform the standard BB84 protocol); (v) applicable to other quantum cryptographic applications (e.g., bit commitment based on the noisy storage model [28]).

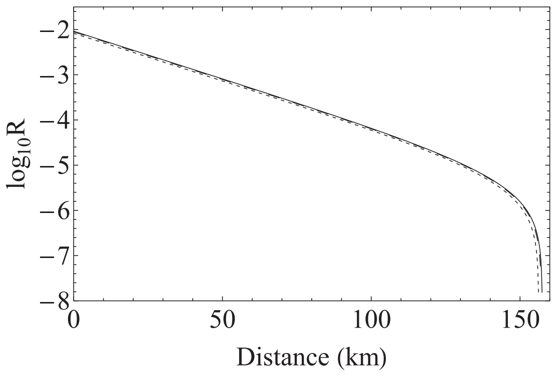

FIG. 2. Lower bound on the secret key rate R of the three-state protocol with decoy states based on our technique. It coincides with that of the decoy-state BB84 protocol (based also on our method) up to the overall factor corresponding to the probability of the basis matched events. For each line, the intensity is optimized, and we use the same experimental parameters and phase modulation error δ as in Fig. 1. Importantly, the three lines almost overlap, i.e., Eve cannot enhance the SPFs by exploiting the channel loss. This shows a dramatic improvement when compared to the results illustrated in Fig. 1.

To simplify the discussion, we assume collective attacks, i.e., Eve applies the same quantum operation to each signal. However, our results also hold against coherent attacks by just applying Azuma's inequality [29–31] (see Appendix A) or the quantum De Finetti theorem [32]. Moreover, we study the asymptotic scenario where the number of signals and decoy pulses is infinite. As the decoy-state method applies to any phase-randomized source, this scenario gives Alice and Bob all the statistics associated with the single-photon part. Also, we assume that Bob's measurement device satisfies two conditions: *random basis choice* and *basis-independent detection efficiency*. The former is fulfilled if Bob selects at random between two or more measurement settings. The latter is satisfied if the detection rate is independent of Bob's choice of the basis. With MDIQKD, we can waive these two conditions and allow the detection system to be untrusted. To illustrate how our general technique works, it is easier to consider a particular example. Specifically, in Sec. II we analyze a prepare-and-measure three-state protocol with SPFs, and in Sec. III we simulate its key generation rate based on our technique and on the GLLP decoy. For comparison, we also plot the key generation rate of this scheme based on the GLLP-SPF and GLLP decoy methods. Next, to see how our method can be applied to MDIQKD, we consider MDIQKD with SPFs in Sec. IV. Finally, we conclude our paper in Sec. V.

II. PREPARE-AND-MEASURE THREE-STATE PROTOCOL

In this section, we consider a prepare-and-measure three-state protocol with SPFs. In the single-photon part of this scheme, Alice is supposed to prepare three states from the Z- and X-basis eigenstates. We assume that she actually prepares imperfect states represented by the density matrices $\hat{\rho}_{0Z}$, $\hat{\rho}_{1Z}$ and $\hat{\rho}_{0X}$, which she selects at random, for instance with probability 1/3, for each signal, and she sends them to Bob. We denote the coefficients of the Bloch vector of $\hat{\rho}_{j\alpha}$, with $j \in \{0,1\}$ and $\alpha \in \{X,Z\}$, as $(P_X^{j\alpha}, P_Y, P_Z^{j\alpha})$. Here, we use the fact that by choosing the Y axis appropriately, we can always set the Y components of all the three Bloch vectors equal to a certain P_Y. Also, we assume that the terminal points of the Bloch vectors form a triangle. That is, all the terminal points cannot be located on a single straight line. On Bob's side, he measures the signals received using either the Z or the X basis, which he selects at random, for instance with probability 1/2, for each incoming signal. After that, Alice and Bob announce their basis choices, and they generate sifted bits and a secure key from those instances where both of them select the Z basis, while the events where Bob chooses the X basis are used for the phase error rate estimation.

Note that this protocol can be regarded as a BB84 scheme with extremely high SPFs since one of the four BB84 states is missing and the state preparation is imperfect. The three states can be mixed, the purification is assumed to be possessed by Alice, and the flaw is independently and identically distributed over the different trials of the state preparation. Moreover, Bob's measurements are not necessarily precise Z and X-basis measurement on a single photon. We denote Bob's positive-operator valued measure (POVM) associated with the basis $\beta \in \{X,Z\}$ as $\{\hat{M}_{0\beta},\hat{M}_{1\beta},\hat{M}_f\}$, where $\hat{M}_f$ and $\hat{M}_{0\beta}$ ($\hat{M}_{1\beta}$) correspond, respectively, to an inconclusive outcome and to the bit value 0 (1). The essential assumption here is that $\hat{M}_f$ is the same for both bases [11].

In the following, we present a precise phase error rate estimation technique. First, we introduce some notation. Let $|\phi_{j\alpha}\rangle_{A_eB}$ denote a purification of $\hat{\rho}_{j\alpha}$, with B and A_e representing, respectively, the system to be sent to Bob and the extended system possessed by Alice. The emission of $\hat{\rho}_{jZ}$ ($\hat{\rho}_{0X}$) is, thus, equivalently expressed by the preparation of $|\Psi_Z\rangle_{AA_eB} = \frac{1}{\sqrt{2}}\sum_{j=0,1}|j_Z\rangle_A|\phi_{jZ}\rangle_{A_eB}$ ($|\Psi_X\rangle_{AA_eB} = |0_X\rangle_A|\phi_{0X}\rangle_{A_eB}$) followed by a Z-basis (X-basis) measurement on system A and then sending only system B. Here, $|j_X\rangle := [|0_Z\rangle + (-1)^j|1_Z\rangle]/\sqrt{2}$. The essential quantity in the proof [11] is the phase error rate e_X, which is a fictitious bit error rate in the X basis, and this is defined as $e_X = (Y^{(Z)}_{0_X,1_X} + Y^{(Z)}_{1_X,0_X})/(Y^{(Z)}_{0_X,0_X} + Y^{(Z)}_{1_X,0_X} + Y^{(Z)}_{0_X,1_X} + Y^{(Z)}_{1_X,1_X})$, where $Y^{(Z)}_{s_\beta,j_\alpha}$, with $s,j \in \{0,1\}$, denotes the joint probability that Alice (Bob) obtains a bit value j (s) given the state preparation of $|\Psi_Z\rangle_{AA_eB}$ and when her (his) basis choice is $\alpha = X$ ($\beta = X$). Importantly, if N_Z denotes the number of reconciled sifted bits, a secure key can be generated by sacrificing $N_Zh(e_X)$ bits in the privacy amplification (PA) [11]. Our technique will give the following theorem.

Theorem. The phase error rate e_X is exactly obtained in the prepare-and-measure three-state protocol.

To see the implication, consider the BB84 scheme with perfect state preparation. From the theorem, we can conclude that even if Alice does not send out one of the BB84 states, the performance remains the same, as the *exact* values of e_X and the bit error rate e_Z can be obtained from the observed data. That is, one of the BB84 states seems to be unnecessary. This means, for instance, that in those implementations of the BB84 protocol that use four laser sources [33] one could keep one laser just as backup in case one of them fails, without any decrease in performance. This also reduces the consumption of random numbers to select the sources. Our security analysis differs from that provided in Ref. [26] in that (i) Ref. [26] considers only a particular set of states and (ii) our analysis requires less PA, which leads to a higher secret key rate. Next, we present the proof of the theorem.

Proof. We first consider the case with $P_Y = 0$. In this case, we can choose $|\phi_{jZ}\rangle_{A_eB}$, the purification of $\hat{\rho}_{jZ}$, and $|\Psi_Z\rangle_{AA_eB}$ as real-valued vectors in the Z basis. To calculate e_X, we consider a virtual protocol where Alice first prepares $|\Psi_Z\rangle_{AA_eB}$ and then both Alice and Bob measure systems A and B in the X basis. In this virtual protocol, Alice sends Bob the virtual (unnormalized) state $\hat{\sigma}_{B;j_X,\text{vir}} = \text{Tr}_{AA_e}[\hat{P}(|j_X\rangle_A) \otimes \hat{\mathbb{1}}_{A_eB}\hat{P}(|\Psi_Z\rangle_{AA_eB})]$, where $\hat{P}(|x\rangle) := |x\rangle\langle x|$ and Tr_{AA_e} represents the partial trace over the systems A and A_e. Since $|\Psi_Z\rangle_{AA_eB}$ is a real-valued vector, the virtual state is also real valued. Therefore, the virtual states lie on the X-Z plane, i.e., they can be expressed as a linear combination of the matrices $\hat{\sigma}_t$, where $\hat{\sigma}_t$, with $t \in \{\text{Id},X,Z\}$, denotes, respectively, the identity and two of the Pauli operators. Therefore, the Bloch vectors of the normalized virtual states are given by $(P_X^{j_X,(\text{vir})},0,P_Z^{j_X,(\text{vir})})$, which are known vectors.

To obtain the terms $Y^{(Z)}_{s_X,j_X}$, we consider its expression $Y^{(Z)}_{s_X,j_X} = \text{Tr}[\hat{\sigma}_{B;j_X,\text{vir}}]\text{Tr}[\hat{D}_{s_X}\hat{\sigma}'_{B;j_X,\text{vir}}]/2$, where $1/2$ is the probability that Bob chooses the X basis, $\text{Tr}[\hat{\sigma}_{\text{B};j_X,\text{vir}}]$ is the probability that the virtual state $\hat{\sigma}_{B;j_X,\text{vir}}$ is emitted, $\hat{\sigma}'_{B;j_X,\text{vir}}$ is the normalized version of $\hat{\sigma}_{B;j_X,\text{vir}}$, and $\hat{D}_{s_X} := \sum_k \hat{A}_k^\dagger\hat{M}_{s_X}\hat{A}_k$ with $\hat{A}_k$ being an arbitrary operator representing Eve's action (see Appendix A). This means that if the transmission rate of $\hat{\sigma}_t$, which we define as $q_{s_X|t} = \text{Tr}(\hat{D}_{s_X}\hat{\sigma}_t)/2$, is obtained, then we can obtain the transmission rate of the virtual states $Y^{(Z)}_{s_X,j_X}$ as $\text{Tr}[\hat{\sigma}_{B;j_X,\text{vir}}](q_{s_X|\text{Id}} + P_X^{j_X,(\text{vir})}q_{s_X|X} + P_Z^{j_X,(\text{vir})}q_{s_X|Z})$. To calculate $q_{s_X|\text{Id}}$, $q_{s_X|X}$, and $q_{s_X|Z}$, recall that the actual three states lie also on the X-Z plane, and we have the following linear equations from the experimentally available data: $(Y^{(Z)}_{s_X,0_Z},Y^{(Z)}_{s_X,1_Z},Y^{(X)}_{s_X,0_X}) = (q_{s_X|\text{Id}},q_{s_X|X},q_{s_X|Z})\hat{A}/6$, where $1/6$ is the joint probability of sending one of the three states and Bob chooses the X basis, and $\hat{A} := (\vec{V}_{0Z}^T,\vec{V}_{1Z}^T,\vec{V}_{0X}^T)$ with $\vec{V}_{j\alpha} := (1,P_X^{j\alpha},P_Z^{j\alpha})$ with T representing the transpose. Since the three Bloch vectors form a triangle, there exists $\hat{A}^{-1}$ to obtain $(q_{s_X|\text{Id}},q_{s_X|X},q_{s_X|Z}) = 6(Y^{(Z)}_{s_X,0_Z},Y^{(Z)}_{s_X,1_Z},Y^{(X)}_{s_X,0_X})\hat{A}^{-1}$. It follows that $Y^{(Z)}_{s_X,j_X}$ can be expressed precisely as a function of $Y^{(Z)}_{s_X,0_Z}$, $Y^{(Z)}_{s_X,1_Z}$, and $Y^{(X)}_{s_X,0_X}$, leading to the exact phase error rate.

The generalization of the above proof to the case where $P_Y \neq 0$ is straightforward. In this case we consider a filter operation whose successful operation is described by $\hat{F} := q|0_Y\rangle\langle 0_Y| + (1-q)|1_Y\rangle\langle 1_Y|$, with $0 \leqslant q \leqslant 1$ and $q \neq 1/2$. That is, this operation uniformly lifts up all the states on the X-Z plane of the Bloch sphere and transforms a state $\hat{\Theta}$, with Bloch vector $\{P_X,0,P_Z\}$, to the state $\hat{F}\hat{\Theta}\hat{F}^\dagger$, whose Bloch vector is $\{f(q)P_X,(2q-1)/(1-2q+2q^2),f(q)P_Z\}$ with $f(q) = 2(1-q)q/(1-2q+2q^2)$. Moreover, the success probability p of the filtering operation is the same for all the states on the X-Z plane, where p is given by $p = q^2 - q + 1/2$. These facts suggest that any normalized qubit state $\hat{\rho}$ can be expressed by $\hat{F}\hat{\rho}_{XZ}\hat{F}^\dagger/p$, where $\hat{\rho}_{XZ}$ is the density operator on the X-Z plane. From this, we can rewrite the transmission rate of the actual states as $\text{Tr}[\hat{\rho}_{XZ}\hat{D}'_{s_X}]$, where $\hat{D}'_{s_X} := \hat{F}^\dagger\hat{D}_{s_X}\hat{F}/p$. This means that even if Alice sends out $\hat{\rho}$ in reality, we are allowed to work on $\hat{\rho}_{XZ}$, that is, we can convert the problem with $P_Y \neq 0$ to the case with $P_Y = 0$. This is so because the proof with $P_Y = 0$ is valid for any matrix $\hat{D}_{s_X}$. This ends the proof. ■

By generalizing the discussion above, it can be seen that when the terminal points of the Bloch vectors of four states form a triangular pyramid (tilted four states), one can obtain the transmission rate of $\hat{\sigma}_Y$ as well, and it follows that the transmission rate of *any* qubit state is also obtained (see Appendix B). In prepare-and-measure QKD-PRS protocols, this implies that when Bob receives qubit states [17] and performs a measurement in the Y basis, then Alice and Bob could use the fictitious bit error rate in the Y basis to improve the secret key rate by applying the data postprocessing of the six-state protocol [17]. Also, note that the qubit assumption could be avoided by using techniques introduced in [34–36].

III. SIMULATION

In this section, we evaluate the performance of a three-state protocol based on phase-randomized weak coherent pulses together with an infinite number of decoy states. As an example of SPFs, we assume modulation errors in the phase-coding scheme; however, our analysis applies as well to any coding scheme. More precisely, we consider that Alice sends Bob signals of the form $|e^{i\xi}\sqrt{\alpha}\rangle_r|e^{i(\xi+\theta_A+\delta\theta_A/\pi)}\sqrt{\alpha}\rangle_s$, where $\xi \in [0,2\pi)$ is a random phase, $\theta_A \in \{0,\pi/2,\pi\}$ encodes Alice's information, the term $\delta\theta_A/\pi$ with $\delta \geqslant 0$ models an instance of phase modulation errors, and $|e^{i\xi}\sqrt{\alpha}\rangle_r$ is a coherent state with mean photon number α. The subscripts r and s are used to denote, respectively, the reference and signal modes. In addition, we assume the same error model on Bob's side, i.e., his phase modulation is $\theta_B - \delta\theta_B/\pi$ when he chooses $\theta_B \in \{0,-\pi/2\}$. Note that Bob does not need to characterize his modulation errors, and since $\delta \geqslant 0$, Alice's and Bob's modulation errors do not cancel each other.

The resulting lower bound on the secret key rate R for different values δ is shown in Fig. 2 (see Appendix C). For comparison, Fig. 1 shows R for the asymptotic decoy-state BB84 protocol with the same phase modulation model with $\theta_A \in \{0,\pi/2,\pi,3\pi/2\}$; its analysis is based on the GLLP-decoy and GLLP-SPF methods [8,10,18,22]. In this protocol, we assume that Alice and Bob choose the bases with probability 1/2 for each pulse. Importantly, since the Bloch vectors of the three states, which are defined in the subspace containing a single photon in modes rs, form a triangle on the X-Z plane, we can obtain the phase error rate precisely. In fact, as shown in Fig. 2, our technique significantly outperforms the GLLP-SPF method in the presence of modulation errors. While the key rate based on the GLLP-SPF method decreases rapidly when δ increases (since it considers the worst-case scenario where losses can increase the fidelity flaws [13]), our method produces an almost constant key rate independently of δ. The slight performance decrease of the three-state protocol is due to the slight increase of the error rates with the increase of δ.

IV. MEASUREMENT-DEVICE-INDEPENDENT QKD

In this section, in order to see that our method is applicable to MDIQKD, we consider the three-state protocol in the MDIQKD setting, i.e., Alice and Bob send Eve the same three states described above. Eve performs a measurement on those states, and then she announces the results. Alice and Bob keep the data associated with the successful results, and Bob applies a bit flip to part of his data [12]. They use the Z basis for key distillation, and the remaining data for parameter estimation.

In the following, we apply our method to MDIQKD. Now, $e_X = (Y^{(Z)}_{\phi^+,0_X1_X} + Y^{(Z)}_{\phi^+,1_X0_X})/(Y^{(Z)}_{\phi^+,0_X1_X} + Y^{(Z)}_{\phi^+,1_X0_X} + Y^{(Z)}_{\phi^+,0_X0_X} + Y^{(Z)}_{\phi^+,1_X1_X})$, where $Y^{(Z)}_{\phi^+,j_Xs_X}$ denotes the joint probability that Alice (Bob) obtains j (s) in the virtual X-basis measurement on system A (B) given the state preparation of $|\Psi_Z\rangle_{AA_eC}$ ($|\Psi_Z\rangle_{BB_eC'}$), and Eve declares an outcome $|\phi^+\rangle$ of her measurement on systems C and C' sent by Alice and Bob. This probability can be expressed as $Y^{(Z)}_{\phi^+,j_Xs_X} = \text{Tr}[\hat{\sigma}_{C;j_X,\text{vir}}]\text{Tr}[\hat{\sigma}_{C';s_X,\text{vir}}]\text{Tr}[\hat{D}_{\phi^+}\hat{\sigma}'_{C;j_X,\text{vir}} \otimes \hat{\sigma}'_{C';s_X,\text{vir}}]$, where $\hat{D}_{\phi^+}$ is Eve's POVM element corresponding to the announcement of the outcome $|\phi^+\rangle$. Here, $\text{Tr}[\hat{\sigma}_{C;j_X,\text{vir}}]$ and $\text{Tr}[\hat{\sigma}_{C';s_X,\text{vir}}]$ are known probabilities, and, therefore, once we have $q_{\phi^+|t,t'} = \text{Tr}(\hat{D}_{\phi^+}\hat{\sigma}_t \otimes \hat{\sigma}_{t'})/4$, with $t' \in \{\text{Id},X,Z\}$, we can obtain the exact value of $Y^{(Z)}_{\phi^+,j_Xs_X}$. This is so because $\hat{\sigma}'_{C;j_X,\text{vir}} \otimes \hat{\sigma}'_{C';s_X,\text{vir}}$ can be expressed as a linear combination of $\hat{\sigma}_t \otimes \hat{\sigma}_{t'}$. Since the experiment gives us nine linear and independent equations with nine unknown parameters $q_{\phi^+|t,t'}$ (see Appendix D), we can obtain $q_{\phi^+|t,t'}$ and the exact phase error rate. By generalizing this idea, we have that the tilted four states provide us with the fictitious bit error rate in the Y basis, which improves the key rate by applying the data postprocessing of the six-state protocol.

V. DISCUSSION AND CONCLUSION

We have analyzed the phase error rate estimation problem that affects the PA step of a QKD-PRS protocol. To generate a secure key, however, it is also important that the bit error rate, which affects the error correction step, is small enough. In this respect, our analysis suggests that while it is important to have a precise state preparation in the key generation basis, that of the other basis is not as essential, which simplifies experimental implementations. For instance, with our results, MDIQKD needs to align only one basis well and can tolerate substantial errors in the alignment of the other bases.

Finally, we would like to emphasize that our technique requires a complete characterization of the signal states that are transmitted [37]. In practice, however, it might be easier to estimate a set of states that very likely contains the signals prepared. In this case, one could directly apply our method by just selecting the signals from that set that minimize the key rate.

To conclude, we have introduced a general phase error rate estimation method that makes the effect of state preparation flaws in phase-randomized sources negligible. We have applied this technique to different QKD protocols and we have shown that it can provide a substantial improvement (in both the achievable rate and distance) when compared to previous results introduced in the GLLP method. Our work constitutes an important step towards secure QKD with imperfect devices.

ACKNOWLEDGMENTS

We thank C.-H. F. Fung, F. Xu, M. Lucamarini, X.-B. Wang, and A. Mizutani for fruitful discussions. We acknowledge support from the National Institute of Information and Communications Technology (NICT) of Japan (Project "Secure Photonic Network Technology" as part of Project UQCC), the Japan Society for the Promotion of Science (JSPS) through its Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program), the European Regional Development Fund (ERDF), the Spanish National Research and Development Program under Project Terasense No. CSD2008-00068 (Consolider-Ingenio), and the Galician Regional Government (Projects No. CN2012/279 and No. CN 2012/260, consolidation of Research Units AtlantTIC and the program "Ayudas para Proyectos de Investigación Desarrollados por Investigadores Emergentes") NSERC, and the CRC program.

APPENDIX A: THREE-STATE PROTOCOL AND COHERENT ATTACKS

Here we present the security proof for the three-state protocol. We consider that Alice and Bob distill key only from those events where both of them use the Z basis, while the events where Bob employs the X basis are used for parameter estimation.

As already introduced in the main text, the preparation of the Z-basis states can be equivalently described as follows. Alice first generates $|\Psi_Z\rangle_{AA_eB} = \frac{1}{\sqrt{2}}\sum_{j=0,1}|j_Z\rangle_A|\phi_{jZ}\rangle_{A_eB}$, and, afterwards, she measures system A in the Z basis, keeps system A_e, and sends Bob system B. We denote this Z-basis measurement as $\hat{Z}_A$. Likewise, the preparation of the signal $|0_X\rangle$ can also be formulated as a two-step process, i.e., Alice first produces $|\Psi_X\rangle_{AA_eB} = |0_X\rangle_A|\phi_{0X}\rangle_{A_eB}$ and then she measures system A in the X basis and sends system B to Bob. This X-basis measurement is denoted as $\hat{X}_A$. To prove the security of the bit values generated from the events where Alice and Bob choose the Z basis given the preparation of $|\Psi_Z\rangle_{AA_eB}$, we need to estimate the phase error rate, which is defined through the virtual events where Alice and Bob choose the X basis rather than the Z basis, given the preparation of $|\Psi_Z\rangle_{AA_eB}$. In these virtual events, Alice sends out $\hat{\sigma}_B^{(1)} := \hat{\sigma}_{B;0_X,\mathrm{vir}}$ and $\hat{\sigma}_B^{(2)} := \hat{\sigma}_{B;1_X,\mathrm{vir}}$. An essential assumption is that the operator $\hat{M}_f$, which corresponds to the failure event of outputting a bit value, is the same for both bases X and Z. Conceptually, this means that Bob could have heralded the receipt of a state from Alice before he decides the measurement basis. This conceptual ability to postpone the measurement basis choice is crucial for the security proof to go through. For the estimation of the phase error rate, we exploit the events where Alice sends out the three actual states, that is, $\hat{\sigma}_B^{(3)} := \hat{\rho}_{0Z}$, $\hat{\sigma}_B^{(4)} := \hat{\rho}_{1Z}$, and $\hat{\sigma}_B^{(5)} := \hat{\rho}_{0X}$, and Bob performs the X-basis measurement. This situation is illustrated in Fig. 3. In this figure, we show Bob's Z-basis measurement to illustrate the actual event, but we do not use this event for the phase error rate estimation.

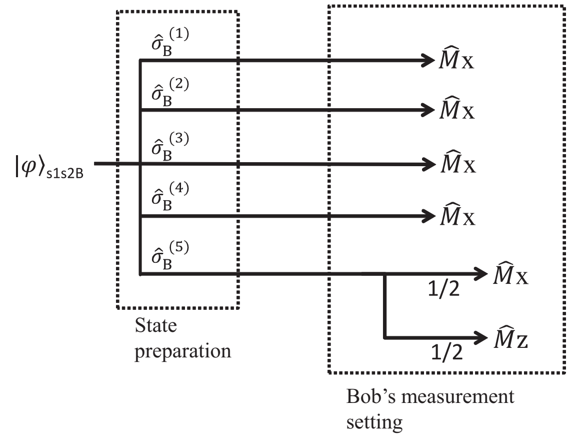

FIG. 3. This diagram illustrates Alice's state preparation process together with Bob's POVM $\hat{M}_X$, the signal state sent by Alice (denoted as "state preparation"), and Bob's measurement setting. The upper two lines correspond to the virtual protocol; the phase error rate e_X is defined in terms of its outcomes. The signals $\hat{\sigma}_B^{(c)}$ with $c = 1,2$ ($c = 3,4,5$) represent the virtual states (the states defined in the protocol). We do not use $\hat{M}_Z$ for the phase error estimation.

In what follows, we consider the probability distribution for the different paths in Fig. 3. The selection of the state that is sent to Bob can be equivalently represented by the preparation of

$$|\varphi\rangle_{s1s2B} := \sum_{c=1}^{5}\sum_{n=0}^{1}\sqrt{P(c)P_c(n)}|c\rangle_{s1}|n\rangle_{s2}|\phi^{(c,n)}\rangle_B\,, \qquad \text{(A1)}$$

followed by an orthogonal measurement using the basis $\{|c\rangle\}$ on the first shield system $s1$ possessed by Alice. Here, $P(c)$ is the probability that Alice sends $\hat{\sigma}_B^{(c)}$, and the second shield system $s2$ represents a system in Alice's laboratory that contains the noise information, i.e., $\sum_{n=0,1}P_c(n)\hat{P}(|\phi^{(c,n)}\rangle) = \hat{\sigma}_B^{(c)}$ with $\hat{P}(|x\rangle) := |x\rangle\langle x|$, where the index $c = 1,2,\ldots,5$ identifies the five possible "sending states" shown in Fig. 3. That is, $\hat{\sigma}_B^{(1)} := \hat{\sigma}_{B;0_X,\mathrm{vir}}$, $\hat{\sigma}_B^{(2)} := \hat{\sigma}_{B;1_X,\mathrm{vir}}$, $\hat{\sigma}_B^{(3)} := \hat{\rho}_{0Z}$, $\hat{\sigma}_B^{(4)} := \hat{\rho}_{1Z}$, and $\hat{\sigma}_B^{(5)} := \hat{\rho}_{0X}$.

Suppose that Alice prepares many systems of the form given by Eq. (A1) and sends system B to Bob through the quantum channel. Also, suppose that Alice and Bob measure in order the shield and B systems using respectively the basis $\{|c\rangle\}$ and the POVM $\hat{M}_X$, and let us consider the lth run of the protocol. According to Azuma's inequality, once we obtain the probabilities for the different paths of Fig. 3 in the lth run conditioned on all previous measurement outcomes, we can determine the actual occurrence number of the corresponding events (see Refs. [30,31] for a proof of this statement).

Next, we calculate these conditional probabilities. For this, let $|\Phi\rangle_{s1s2B} = |\varphi_{l-1}\rangle_{s1s2B}|\varphi_l\rangle_{s1s2B}|\varphi_r\rangle_{s1s2B}$ denote the state prepared by Alice in an execution of the protocol. Here, $|\varphi_{l-1}\rangle_{s1s2B}$, $|\varphi_l\rangle_{s1s2B}$, and $|\varphi_r\rangle_{s1s2B}$ represent, respectively, Alice's signals in the first $l-1$ runs, in the lth run, and in the rest of runs.

This state evolves according to Eve's unitary transformation $\hat{V}_{\mathrm{BE}}$ on Bob's system B and on her system E as follows:

$$\hat{V}_{\mathrm{BE}}|\Phi\rangle_{s1s2B}|0\rangle_E = \sum_k \hat{B}_{kB}|\Phi\rangle_{s1s2B}|k\rangle_E, \qquad \text{(A2)}$$

where $\hat{B}_{kB}$ is a Kraus operator acting on system B. Importantly, $\hat{V}_{\mathrm{BE}}$ and $\hat{B}_{kB}$ are independent of the state preparation process. This is so because the classical communication between Alice and Bob is done *after* finishing the measurements. Let the joint operator

$$\hat{O}_{l-1,s1B} = \otimes_{u=1}^{l-1}\hat{M}_{s1_uB_u}, \qquad \text{(A3)}$$

where $\hat{M}_{s1_u,s_u}$ denotes the Kraus operator associated with the uth measurement outcome of the first shield system $s1$ and Bob's uth measurement outcome B. The normalized state of the lth system $s1B$ conditioned on the measurement outcomes O_{l-1} of the first $l-1$ joint systems, then becomes

$$\begin{aligned}
\hat{\rho}^B_{l|O_{l-1}} &= \hat{\sigma}^B_{l|O_{l-1}}/p^{(l)}\,,\\
\hat{\sigma}^B_{l|O_{l-1}} &:= \sum_n\sum_{c,c'}\sqrt{P(c)P_c(n)P(c')P_{c'}(n)}\\
&\quad\times\sum_k\sum_{\vec{x}_{l-1},\vec{x}_r}\hat{A}^{(\vec{x}_{l-1},\vec{x}_r)}_{k,B|O_{l-1}}|\phi^{(c,n)}\rangle_B\langle\phi^{(c',n)}|\hat{A}^{\dagger(\vec{x}_{l-1},\vec{x}_r)}_{k,B|O_{l-1}},\\
p^{(l)} &: = \mathrm{Tr}\big(\hat{\sigma}^B_{l_c|O_{l-1}}\big),
\end{aligned} \qquad \text{(A4)}$$

where

$$\hat{A}^{(\vec{x}_{l-1},\vec{x}_r)}_{k,B|O_{l-1}} := \langle\vec{x}_r|\langle\vec{x}_{l-1}|\hat{O}_{l-1,s1s2B}\hat{B}_{kB}|\varphi_{l-1}\rangle_{s1s2B}|\varphi_r\rangle_{s1s2B}.$$

Here $\{\langle\vec{x}_r|\}$ ($\{\langle\vec{x}_{l-1}|\}$) represents a basis for all the systems $s1s2B$ after the lth run (for all the systems $s1s2B$ for the first $l-1$ runs). Importantly, Eq. (A4) states that the lth joint system is subjected to Eve's action and her action depends on all the previous measurement outcomes on the first $l-1$ joint systems.

Now, the probability of obtaining c and the bit value s_X conditioned on O_{l-1} is given by

$$\begin{aligned}P_{s_X,c|O_{l-1}} &= \frac{P(c\wedge X)}{p^{(l)}}\sum_k\sum_{\vec{x}_{l-1},\vec{x}_r}\mathrm{Tr}\big[\hat{A}^{(\vec{x}_{l-1},\vec{x}_r)}_{k,B|O_{l-1}}\hat{\sigma}^{(c)}_B\\ &\quad\times\hat{A}^{\dagger(\vec{x}_{l-1},\vec{x}_r)}_{k,B|O_{l-1}}\hat{M}_{s_X}\big]\\ &:= \frac{P(c\wedge X)}{p^{(l)}}\mathrm{Tr}\big[\hat{D}_{s_X|O_{l-1}}\hat{\sigma}^{(c)}_B\big],\end{aligned} \tag{A5}$$

where we use $\sum_{n=0,1}P_c(n)\hat{P}(|\phi^{(c,n)}\rangle)=\hat{\sigma}^{(c)}_B$, $P(c\wedge X)$ is the joint probability that Alice obtains c and Bob chooses the X basis, and $\hat{D}_{s_X|O_{l-1}}$ is defined by

$$\hat{D}_{s_X|O_{l-1}} = \sum_k\sum_{\vec{x}_{l-1},\vec{x}_r}\hat{A}^{\dagger(\vec{x}_{l-1},\vec{x}_r)}_{k,B|O_{l-1}}\hat{M}_{s_X}\hat{A}^{(\vec{x}_{l-1},\vec{x}_r)}_{k,B|O_{l-1}}.$$

Notice that $P_{s_X,c|O_{l-1}}$, with $c=1,2$ ($c=3,4,5$), is the same as $P(Z)Y^{(Z)}_{s_X,j_X}$ ($P(\alpha)Y^{(\alpha)}_{s_X,j_\alpha}$) in the main text except that $P_{s_X,c|O_{l-1}}$ is the conditional probability, where $P(Z)$ ($P(\alpha)$) is the probability that Alice chooses the Z (α) basis. Also, note that the discussions in the main text use the probability $P(c\wedge X)$ given in Eq. (A5) but do not employ the explicit form of $\hat{D}_{s_X}$. Therefore, the relationships among $Y^{(Z)}_{s_X,j_X}$, $Y^{(Z)}_{s_X,0_Z}$, $Y^{(Z)}_{s_X,1_Z}$, and $Y^{(X)}_{s_X,0_X}$ in the main text can be interpreted as linear relationships between the lth *conditional* probabilities $P_{s_X,c|O_{l-1}}$. This is so because the normalization factor $p^{(l)}$ does not affect this interrelation. Thus, by taking the summation of such probabilities over l, Azuma's inequality [30,31] gives the actual occurrence number of such events and the phase error rate in the virtual protocol can be estimated. This concludes the proof.

APPENDIX B: TILTED FOUR-STATE PROTOCOL

In this section we apply our phase error rate estimation technique to a tilted four-state protocol, which is a variant of the BB84 scheme. Suppose that Alice sends Bob four states given by

$$\hat{\rho}_{j\alpha} = \frac{1}{2}\left(\hat{\mathbb{1}}+\sum_{t=x,y,z}P^{j\alpha}_t\hat{\sigma}_t\right), \tag{B1}$$

where $j\in\{0,1\}$ and $\alpha\in\{X,Z\}$. Moreover, let us assume that the vectors $\vec{V}_{j\alpha}$, with $\vec{V}_{j\alpha}=(1,P^{j\alpha}_X,P^{j\alpha}_Y,P^{j\alpha}_Z)$, are mutually linearly independent. From the viewpoint of the Bloch sphere, this means that the terminal points of the four Bloch vectors associated with the four states form a triangular pyramid. Suppose also that Alice and Bob distill key from the Z basis and use the events where Bob employs the X basis for parameter estimation. In this scenario, let $|\phi_{jZ}\rangle_{A_e,B}$ denote a purification of $\hat{\rho}_{jZ}$, with A_e and B representing, respectively, Alice's shield system and the system that is sent to Bob. With this notation, Alice's state preparation process in the Z basis can be described by using either of the following two source states:

$$|\Psi_Z\rangle_{AA_eB} = \frac{1}{\sqrt{2}}\sum_{j=0,1}|j_Z\rangle_A|\phi_{jZ}\rangle_{A_eB}, \tag{B2}$$

$$|\Psi_Z\rangle_{AA_eB} = \frac{1}{\sqrt{2}}\sum_{j=0,1}|j_Z\rangle_A|\phi_{(j\oplus1)Z}\rangle_{A_eB}, \tag{B3}$$

where A is a virtual qubit system of Alice and the symbol $\oplus$ denotes the modulo-2 addition. If Alice measures system A in the Z basis she prepares the desired state at site B, while she keeps the shield system A_e. Here, the difference between Eqs. (B2) and (B3) is just a bit flip. Since a bit flip is a symmetry in the problem, Alice is allowed to choose either of the two equations above [Eqs. (B2) and (B3)] in constructing the purifications with the goal of optimizing the key generation rate.

To calculate the phase error rate e_X we consider the virtual protocol where Alice and Bob measure $|\Psi_Z\rangle_{AA_eB}$ in the X basis. In this virtual protocol, Alice emits

$$\hat{\sigma}_{B;j_X,\mathrm{vir}} = \mathrm{Tr}_{A,A_e}[\hat{P}(|j_X\rangle_A)\otimes\hat{\mathbb{1}}_{A_eB}\hat{P}(|\Psi_Z\rangle_{AA_eB})].$$

We denote these signals $\hat{\sigma}_{B;j_X,\mathrm{vir}}$ as virtual states, and we define the normalized state $\hat{\sigma}'_{B;j_X,\mathrm{vir}}=\hat{\sigma}_{B;j_X,\mathrm{vir}}/\mathrm{Tr}(\hat{\sigma}_{B;j_X,\mathrm{vir}})$. Then the joint probability that Alice sends $\hat{\sigma}_{B;j_X,\mathrm{vir}}$ and Bob detects the bit value s_X given the state preparation of $|\Psi_Z\rangle_{AA_eB}$ is expressed by $Y^{(\mathrm{Z})}_{s_X,j_X}=\mathrm{Tr}[\hat{\sigma}_{B;j_X,\mathrm{vir}}]\mathrm{Tr}(\hat{D}_{s_X}\hat{\sigma}'_{B;j_X,\mathrm{vir}})/2$. Note that the value of $\mathrm{Tr}[\hat{\sigma}_{B;j_X,\mathrm{vir}}]$ is known from the protocol as well as Eqs. (B2) and (B3). The phase error rate e_X is given by

$$e_X = \frac{Y^{(\mathrm{Z})}_{0_X,1_X}+Y^{(\mathrm{Z})}_{1_X,0_X}}{Y^{(\mathrm{Z})}_{0_X,0_X}+Y^{(\mathrm{Z})}_{1_X,0_X}+Y^{(\mathrm{Z})}_{0_X,1_X}+Y^{(\mathrm{Z})}_{1_X,1_X}}. \tag{B4}$$

Now, since $\hat{\sigma}'_{B;j_X,\mathrm{vir}}$ can also be written as

$$\hat{\sigma}'_{B;j_X,\mathrm{vir}} = \frac{1}{2}\left(\hat{\mathbb{1}}+\sum_{t=X,Y,Z}P^{j_X,(\mathrm{vir})}_t\hat{\sigma}_t\right), \tag{B5}$$

where $P^{j_X,(\mathrm{vir})}_t$ is the t component of the Bloch vector, in order to obtain $Y^{(\mathrm{Z})}_{s_X,j_X}$ (and thus e_X) it is enough to calculate $q_{s_X|t}=\mathrm{Tr}(\hat{D}_{s_X}\hat{\sigma}_t)/2$ with $t\in\{\mathrm{Id},X,Y,Z\}$. For this, note that in the actual experiment we have the following constraints:

$$\begin{aligned}Y^{(\alpha)}_{s_X,j_\alpha} &= P(j_\alpha)\mathrm{Tr}\big(\hat{D}_{s_X}\hat{\rho}_{j\alpha}\big)/2\\ &= P(j_\alpha)\big(q_{s_X|\mathrm{Id}}+P^{j\alpha}_X q_{s_X|x}+P^{j\alpha}_Y q_{s_X|y}+P^{j\alpha}_Z q_{s_X|z}\big)/2.\end{aligned} \tag{B6}$$

Here, $P(j_\alpha)$ is the probability that Alice sends $\hat{\rho}_{j\alpha}$. Then, as long as the vectors $\vec{V}_{j\alpha}$ are mutually linearly independent, we can solve the set of linear equations given by Eq. (B6) and obtain the parameters $q_{s_X|\mathrm{Id}}$, $q_{s_X|x}$, $q_{s_X|y}$, and $q_{s_X|z}$. That is, we can determine the exact transmission rate of *any* state, including the signal $\hat{\sigma}_{B;j_X,\mathrm{vir}}$, which gives the phase error rate.

Moreover, if Bob's POVM elements act on a qubit space and he performs a measurement in the Y basis, Alice and Bob can also estimate the Y-basis error rate. To see this, note that in the discussion above one needs to change only the terms s_X with s_Y and define the Y-basis virtual state as

$$\hat{\sigma}_{B;j_Y,\mathrm{vir}} = \mathrm{Tr}_{AA_e}[\hat{P}(|j_Y\rangle_A) \otimes \hat{\mathbb{1}}_{A_eB}\hat{P}(|\Psi_Z\rangle_{AA_eB})].$$

By following the same argument as in the main text, it is easy to see that the use of the tilted four states in MDIQKD permits estimation of the fictitious Y-basis error rate there also.

APPENDIX C: SIMULATION FOR THE THREE-STATE PROTOCOL

In this section we present the calculations used to obtain Fig. 2 in the main text. We begin with the single-photon components of the signals sent by Alice. In particular, we have that the single-photon part of $|e^{i\xi}\sqrt{\alpha}\rangle_r|e^{i(\xi+\theta_A+\delta\theta_A/\pi)}\sqrt{\alpha}\rangle_s$ is given by $(|1\rangle_r|0\rangle_s + e^{i(\theta_A+\delta\theta_A/\pi)}|0\rangle_r|1\rangle_s)/\sqrt{2}$, where 0 and 1 represent, respectively, the photon number. The set of two pure states $\{|1\rangle_r|0\rangle_s, |0\rangle_r|1\rangle_s\}$ forms a qubit basis. Therefore, we can choose the Z basis such that Alice's Z-basis states and X-basis state are expressed as

$$\begin{aligned}
|\phi_{0Z}\rangle &= |0_Z\rangle, \\
|\phi_{1Z}\rangle &= -\sin\frac{\delta}{2}|0_Z\rangle + \cos\frac{\delta}{2}|1_Z\rangle, \qquad \text{(C1)} \\
|\phi_{0X}\rangle &= \cos\left(\frac{\pi}{4}+\frac{\delta}{4}\right)|0_Z\rangle + \sin\left(\frac{\pi}{4}+\frac{\delta}{4}\right)|1_Z\rangle,
\end{aligned}$$

where $|0_Z\rangle := (|0_Y\rangle + |1_Y\rangle)/\sqrt{2}$ and $|1_Z\rangle := (-i|0_Y\rangle + i|1_Y\rangle)/\sqrt{2}$ with $|0_Y\rangle := |1\rangle_r|0\rangle_s$ and $|1_Y\rangle := |0\rangle_r|1\rangle_s$. In order to discuss the security, we need to consider the virtual protocol where Alice generates the state $(|0_Z\rangle_A|\phi_{0Z}\rangle_B + |1_Z\rangle_A|\phi_{1Z}\rangle_B)/\sqrt{2}$ and sends system B to Bob. From this, one can see that Alice sends out the virtual state $|\phi'_{jX}\rangle_B$ with probability $[1-(-1)^j\sin\frac{\delta}{2}]/2$. Here, the Bloch vector of $|\phi'_{jX}\rangle_B$ is $\vec{V}_{\phi'_{jX}} := [(-1)^j\cos\frac{\delta}{2}, -(-1)^j\sin\frac{\delta}{2}]$ where the first (second) element corresponds to the X (Z) component of the Bloch vector.

For our simulations, we consider a channel model where the conditional probabilities $V_{s_X||\theta\rangle}$ [i.e., the conditional probability that Bob obtains s_X given that Alice sent him the state $|\theta\rangle := \cos\theta/2|0_Z\rangle + \sin\theta/2|1_Z\rangle$] are given by

$$\begin{aligned}
V_{s_X||\theta\rangle} &= (1-L)C_{s,\theta}(1-e_d) + Le_{\mathrm{d}}(1-e_d) \\
&\quad + \frac{1}{2}\big[(1-L)e_d + Le_d^2\big], \qquad \text{(C2)}
\end{aligned}$$

where $C_{s,\theta} := \mathrm{Tr}[\hat{P}(|\theta-\pi/2+\delta/2\rangle)\hat{P}(|s_Z\rangle)]$, e_d is the dark count rate of Bob's detectors, and L denotes the total loss rate of the channel, Bob's single-photon detectors, and the 50% loss of Bob's Mach-Zehnder interferometer. Here, $-\pi/2+\delta/2$ represents Bob's phase modulation with an imperfect phase modulator, and the first (second) term models a single detection click at Bob's side produced by a photon (dark count), while the last term represents simultaneous clicks. Note that in this last case (simultaneous clicks), Bob assigns a random bit value to the measurement result. From $V_{s_X||\theta\rangle}$ as well as the fact that Alice sends out all the three states with probability 1/3, we can readily obtain the experimental data $Y^{(Z)}_{s_X,j_Z}$ and $Y^{(X)}_{s_X,0_X}$. By applying the technique introduced in the main text, we can also obtain the precise phase error rate. Note that by using the basis-independent detection efficiency the single-photon gain $Q_Z^{(1)}$ is given by

$$Q_Z^{(1)} = \frac{1}{3}e^{-2\alpha}\alpha\sum_{s,j}V_{s_X||\phi_{jZ}\rangle}/2, \qquad \text{(C3)}$$

where $\frac{1}{3}$ is the probability that both of Alice and Bob choose the Z basis, and 1/2 is the probability that Alice sends $|\phi_{jZ}\rangle$ given she chooses the Z basis.

Similarly, one can obtain the overall gain Q_Z and the bit error rate e_Z in the Z basis. These parameters have the forms

$$\begin{aligned}
Q_Z &= \tfrac{1}{3}[P_{0|0}(1-P_{1|0}) + (1-P_{0|0})P_{1|0} + P_{0|0}P_{1|0}] \\
&\quad + \tfrac{1}{3}[P_{0|1}(1-P_{1|1}) + (1-P_{0|1})P_{1|1} + P_{0|1}P_{1|1}],
\end{aligned} \qquad \text{(C4)}$$

$$\begin{aligned}
w_Z &= \tfrac{1}{3}[P_{1|0}(1-P_{0|0}) + P_{1|0}P_{0|0}/2 + P_{0|1}(1-P_{1|1}) \\
&\quad + P_{0|1}P_{1|1}/2], \\
e_Z &= w_Z/Q_Z, \qquad \text{(C5)}
\end{aligned}$$

where $P_{s|j}$ is the conditional probability that Bob obtains a bit value s given that Alice sent him a bit value j. These probabilities can be written as

$$P_{0|0} = e_d + (1-e_d)[1-e^{-\alpha(1-L)}], \qquad \text{(C6)}$$

$$P_{1|0} = e_d, \qquad \text{(C7)}$$

$$P_{0|1} = e_d + (1-e_{\mathrm{d}})\big[1-e^{-\alpha(1-L)\sin^2(\delta/2)}\big], \qquad \text{(C8)}$$

$$P_{1|1} = e_d + (1-e_d)\big[1-e^{-\alpha(1-L)\cos^2(\delta/2)}\big]. \qquad \text{(C9)}$$

Finally, the asymptotic key generation rate is given by

$$R = Q_Z^{(1)}[1-h(e_X)] - Q_Zh(e_Z), \qquad \text{(C10)}$$

where $h(x)$ is the binary entropy function. For each value of the distance, we optimize the parameter α to maximize the key rate. The result is shown in Fig. 2 in the main text.

APPENDIX D: THREE-STATE MDIQKD

Here, we present the set of linear equations that are obtained from an MDIQKD experiment based on the three-state protocol, which are expressed as

$$Y^{(ZZ)}_{\phi^+,j_Zs_Z} = \frac{\gamma}{9}\mathrm{Tr}[\hat{D}_{\phi^+}\hat{\rho}_{jZ}\otimes\hat{\rho}_{sZ}], \qquad \text{(D1)}$$

$$Y^{(XZ)}_{\phi^+,0_Xs_Z} = \tfrac{1}{9}\mathrm{Tr}[\hat{D}_{\phi^+}\hat{\rho}_{0X}\otimes\hat{\rho}_{sZ}], \qquad \text{(D2)}$$

$$Y^{(ZX)}_{\phi^+,j_Z0_X} = \tfrac{1}{9}\mathrm{Tr}[\hat{D}_{\phi^+}\hat{\rho}_{jZ}\otimes\hat{\rho}_{0X}], \qquad \text{(D3)}$$

$$Y^{(XX)}_{\phi^+,0_X0_X} = \tfrac{1}{9}\mathrm{Tr}[\hat{D}_{\phi^+}\hat{\rho}_{0X}\otimes\hat{\rho}_{0X}], \qquad \text{(D4)}$$

where $\hat{\rho}_{j\alpha} = \frac{1}{2}(\hat{\mathbb{1}} + P_X^{j\alpha}\hat{\sigma}_X + P_Z^{j\alpha}\hat{\sigma}_Z)$ and $\frac{\gamma}{9}$ (with $0 < \gamma < 1$) is the probability that Alice and Bob send Charles $\hat{\rho}_{jZ}$ and $\hat{\rho}_{sZ}$, respectively, and they sacrifice such instances as test bits. Note that unlike the prepare-and-measure three-state protocol, Alice and Bob have to sacrifice the test bits for the estimation of the phase error rate, and they generate a key from the rest of the instances where they choose the Z basis and Charles declares ϕ^+. Importantly, these equations are independent of each other as we have assumed that the terminal points of the Bloch vectors of the three states form a triangle. Therefore, we can precisely obtain $q_{\phi^+|t,t'}$ and the exact phase error rate.

[1] M. Dušek, N. Lütkenhaus and M. Hendrych, Prog. Opt. **49**, 381 (2006); D. Roar Hjelme, L. Lydersen, and V. Makarov, arXiv:1108.1718.

[2] M. Sasaki *et al.*, Opt. Express **19**, 10387 (2011).

[3] K. Yoshino, T. Ochi, M. Fujiwara, M. Sasaki, and A. Tajima, Opt. Express **21**, 31395 (2013).

[4] D. Mayers, in *Advances in Cryptography—Proceedings of Crypto96* (Springer, New York, 1996), pp. 343–357; J. Assoc. Comput. Mach. **48**, 351 (2001).

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

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

[7] M. Koashi and J. Preskill, Phys. Rev. Lett. **90**, 057902 (2003).

[8] M. Koashi, New J. Phys. **11**, 045018 (2009).

[9] B. Kraus, N. Gisin, and R. Renner, Phys. Rev. Lett. **95**, 080501 (2005); R. Renner, Ph.D. dissertation No. 16242, ETH 2005, arXiv:quant-ph/0512258.

[10] M. Tomamichel and R. Renner, Phys. Rev. Lett. **106**, 110506 (2011); M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, Nat. Commun. **3**, 634 (2012).

[11] M. Koashi, arXiv:0704.3661.

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

[13] K. Tamaki, H.-K. Lo, C.-H. F. Fung, and B. Qi, Phys. Rev. A **86**, 059903(E) (2012).

[14] X. Ma, C-H. F. Fung, and M. Razavi, Phys. Rev. A **86**, 052305 (2012); X.-B. Wang, *ibid.* **87**, 012320 (2013); F. Xu, M. Curty, B. Qi, and H.-K. Lo, New J. Phys. **15**, 113007 (2013); M. Curty, F. Xu, W. Cui, C. C. W. Lim, K. Tamaki, and H.-K. Lo, Nat. Commun. **5**, 3732 (2014); A. Mizutani, K. Tamaki, R. Ikuta, T. Yamamoto, and N. Imoto, Sci. Rep. **4**, 5236 (2014).

[15] T. Ferreira da Silva, D. Vitoreti, G. B. Xavier, G. C. do Amaral, G. P. Temporao, and J. P. von der Weid, Phys. Rev. A **88**, 052303 (2013); A. Rubenok, J. A. Slater, P. Chan, I. Lucio-Martinez, and W. Tittel, Phys. Rev. Lett. **111**, 130501 (2013); Y. Liu *et al.*, *ibid.* **111**, 130502 (2013); F. Xu, B. Qi, Z. Liao, and H.-K. Lo, Appl. Phys. Lett. **103**, 061101 (2013).

[16] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India*, IEEE Press (IEEE Computer Society Press, Los Alamitos, CA, 1984), pp. 175–179.

[17] D. Bruss, Phys. Rev. Lett. **81**, 3018 (1998); H.-K. Lo, Quantum Inf. Comput. **1**, 81 (2001).

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

[19] W.-Y. Hwang, Phys. Rev. Lett. **91**, 057901 (2003); H.-K. Lo, X. Ma, and K. Chen, *ibid.* **94**, 230504 (2005); X.-B. Wang, *ibid.* **94**, 230503 (2005).

[20] The dark count rate of Bob's detectors is 0.5×10^{-7}, the overall transmittance of his detection apparatus is 0.15, the loss coefficient of the channel is 0.21 dB/km, and the efficiency of the error correction protocol is 1.22.

[21] T. Honjo, K. Inoue, and H. Takahashi, Opt. Lett. **29**, 2797 (2004); G. Li, Adv. Opt. Photon. **1**, 279 (2009).

[22] Ø. Marøy, L. Lydersen, and J. Skaar, Phys. Rev. A **82**, 032337 (2010).

[23] Note that the work in [22] also takes both modulation and detector flaws into account, but unfortunately it is not loss tolerant.

[24] S. M. Barnett, B. Huttner, and S. J. D. Phoenix, J. Mod. Opt. **40**, 2501 (1993); R. Matsumoto, and S. Watanabe, IEICE Trans. Fundamentals E **91**, 2870 (2008).

[25] S. N. Molotkov and S. S. Nazin, J. Exp. Theor. Phys. **63**, 924 (1996); S. N. Molotkov, *ibid.* **87**, 288 (1998); B.-S. Shi, Y.-K. Jiang, and G.-C. Guo, Appl. Phys. B **70**, 415 (2000).

[26] C.-H. F. Fung and H.-K. Lo, Phys. Rev. A **74**, 042342 (2006).

[27] In contrast to our work, note that the recent results introduced by X.-B. Wang, Phys. Rev. A **87**, 012320 (2013) consider four signals and, even then require that their states are prepared rather precisely.

[28] S. Wehner, M. Curty, C. Schaffner, and H.-K. Lo, Phys. Rev. A **81**, 052336 (2010).

[29] K. Azuma, Tohoku Math. J. **19**, 357 (1967).

[30] J.-C. Boileau, K. Tamaki, J. Batuwantudawe, R. Laflamme, and J. M. Renes, Phys. Rev. Lett. **94**, 040503 (2005).

[31] K. Tamaki, N. Lütkenhaus, M. Koashi, and J. Batuwantudawe, Phys. Rev. A **80**, 032302 (2009).

[32] C. M. Caves and C. A. Fuchs, J. Math. Phys. **43**, 4537 (2002); R. Koenig and R. Renner, *ibid.* **46**, 122108 (2005).

[33] R. J. Hughes, J. E. Nordholt, D. Derkacs, and C. G. Peterson, New. J. Phys. **4**, 43 (2002).

[34] G. Kato and K. Tamaki, arXiv:1008.4663.

[35] C.-H. F. Fung, H. F. Chau, and H.-K. Lo, Phys. Rev. A **84**, 020303(R) (2011).

[36] T. Moroder, M. Curty, and N. Lütkenhaus, New J. Phys. **11**, 045008 (2009).

[37] Note that the recent work by Z.-Q. Yin *et al.*, Phys. Rev. A **88**, 062322 (2013), does not require such complete state characterization, but it assumes perfect single-photon sources, which unfortunately are still unavailable.