Protocol choice and parameter optimization in decoy-state measurement-device-independent quantum key distribution

3 authors:

Feihu Xu
University of Science and Technology of China
187 PUBLICATIONS   **11,658** CITATIONS

He Xu
University of Toronto
6 PUBLICATIONS   **744** CITATIONS

Hoi-Kwong Lo
University of Toronto
252 PUBLICATIONS   **28,874** CITATIONS

Protocol choice and parameter optimization in decoy-state measurement-device-independent quantum key distribution

Feihu Xu[1],[*] He Xu[1], and Hoi-Kwong Lo[1]

[1] Centre for Quantum Information and Quantum Control,
Department of Physics and Department of Electrical & Computer Engineering,
University of Toronto, Toronto, Ontario, M5S 3G4, Canada
(Dated: June 5, 2014)

Measurement-device-independent quantum key distribution (MDI-QKD) has been demonstrated in both laboratories and field-tests using attenuated lasers combined with the decoy-state technique. Although researchers have studied various decoy-state MDI-QKD protocols with two or three decoy states, a clear comparison between these protocols is still missing. This invokes the question of how many types of decoy states are needed for practical MDI-QKD. Moreover, the system parameters to implement decoy-state MDI-QKD are only partially optimized in all previous works, which casts doubt on the actual performance of former demonstrations. Here, we present analytical and numerical decoy-state methods with one, two and three decoy states. We provide a clear comparison among these methods and find that *two* decoy states already enable a near optimal estimation and more decoy states cannot improve the key rate much in either asymptotic or finite-data settings. Furthermore, we perform a full optimization of system parameters and show that full optimization can significantly improve the key rate in the finite-data setting. By simulating a real experiment, we find that full optimization can increase the key rate by more than one order of magnitude compared to non-optimization. A local search method to optimize efficiently the system parameters is proposed. This method can be four orders of magnitude faster than a trivial exhaustive search to achieve a similar optimal key rate. We expect that this local search method could be valuable for general fields in physics.

I. INTRODUCTION

Quantum cryptography or quantum key distribution (QKD) can provide information-theoretic security based on the laws of quantum physics [1, 2]. During the past decade, commercial QKD products have appeared on the market; various field-test QKD networks have already been built around the world [3, 4]. However, owing to the imperfections in real-life implementations of QKD, a gap between its theory and practice remains unfilled. In particular, an eavesdropper (Eve) may exploit these imperfections and launch attacks not covered by the original security proofs of QKD. Indeed, the recent demonstrations of various attacks [5–10] on top of practical QKD systems highlight that this gap is a major problem for the real-life security of QKD.

Fortunately, measurement-device-independent quantum key distribution (MDI-QKD) [11] removes all detector side-channel attacks, the most important security loophole in conventional QKD implementations [5–10]. The key idea of MDI-QKD is that both legitimate users (Alice and Bob) are senders. They transmit signals to an *untrusted* third party, Charles (or Eve), who is supposed to perform a Bell state measurement. Such a measurement provides post-selected entanglement that can be verified by Alice and Bob. By using a decoy-state protocol [12–22], Alice and Bob can use imperfect single-photon sources such as attenuated lasers and still estimate the contributions from single-photon signals. Unlike security patches [23] and (full) device-independent QKD [24–26], MDI-QKD can remove all detector side channels and is also practical for current technology. Hence, it has attracted a lot of scientific attention from the research community on both theoretical [27–33] and experimental [34–37] studies.

Various decoy-state methods have been proposed for MDI-QKD. Ref. [27] proposed a numerical method using two and three decoy states. Refs. [28–32] discussed different analytical approaches based on two decoy states. Experimentally, Refs. [34, 36, 37] implemented the two decoy-state protocol, while Ref. [35] chose the three decoy-state protocol. Consequently, the first open question is: how many types of decoy states are essentially needed for MDI-QKD in practice?

On the other hand, to implement MDI-QKD, one has to know the parameters to optimize the system performance. However, some previous theoretical studies [27–29] and experimental implementations [35, 36] simply choose empirical parameters *without* optimization. Hence, an important question is: how can one optimize the parameters used in MDI-QKD? This question is non-trivial, given the large number of parameters involved. Another question is: how much will a careful parameter optimization improve performance?

It is well known that an efficient version of decoy-state BB84 with biased basis choice rather than the standard one can significantly improve the key rate [20–22]. In biased basis choice, the basis-sift factor can be 1 instead of 1/2 for the standard one, hence the maximal improvement is about 100% [38]. In MDI-QKD, the X basis cannot be used to generate secure keys due to its large error rate [11], thus the basis-sift factor can be improved from 1/4 (standard one) to 1 (biased basis choice). An efficient protocol with biased basis choice has a *larger* improvement up to 300%.

The parameters to implement this efficient protocol are chosen via optimization. Previously, this optimization has been studied and implemented on decoy-state BB84 [15–22] as well as decoy-state MDI-QKD [30–34, 37]. Nonetheless, except for Refs. [20, 22], it is only partial. That is, the basis choice is independent of the intensity choice, which we will call *simplified choice* or partial optimization in this paper. This choice is simple for implementation as the sender's two modulators on the intensity and bit information can be completely

independent. However, from the theoretical point of view, the simplified choice cannot result in the optimal key rate due to the finite-data effect. The *optimal choice* should select the majority of signal state in the Z-basis for key generation, while the majority of decoy states in the X-basis for a good estimation on the phase error rate, *i.e.*, the basis choice should depend on the intensity choice.

In the asymptotic case with infinite data-set, the optimal choice and the simplified choice are the same. In BB84, the optimal choice cannot improve the key rate much, because: a) it is relatively easy to generate a large amount of detection counts (approaching asymptotic case) using a high-speed system [21]; b) the receiver (Bob) cannot implement the optimal choice as he cannot distinguish the signal state from decoy sates. In MDI-QKD, we note however that the optimal choice can significantly increase the key rate, because: a) the detection counts are relatively low[1], *i.e.*, they are away from the asymptotic case; b) both Alice and Bob are the sender and can explicitly know the intensity choice. Indeed, by simulating a real experiment, we find that optimal choice can improve simplified choice about 200% in a reasonable data-set (see Table IV).

Implementing optimal choice requires a *full* parameter optimization with the numerical search over many dimensions including the intensity choice of signal state and decoy states and the probability choice of intensities and bases. With a trivial exhaustive search, such an optimization problem is believed to be a challenge in terms of computational complexity (see Table I as well as the case in decoy-state BB84 [22]). This might be one of the major reasons that a full parameter optimization is neglected in all previous works on decoy-state MDI-QKD [27–37]. Hence, another open question is: how one can perform a full parameter optimization in MDI-QKD?

In this paper, we provide solutions to the above open questions. We present analytical and numerical decoy-state methods with one, two and three decoy states. By clearly comparing these methods, we find that two decoy states combined with a full parameter optimization is already close to the optimal estimation and more decoy states cannot improve the key rate much in both asymptotic and finite-data cases. Moreover, we introduce a local search algorithm (LSA) [39], a well-known algorithm in the field of computer science, to QKD for a full parameter optimization. This algorithm requires very low computational power and can be four orders of magnitude faster than a trivial exhaustive search to achieve a similar optimal key rate. It can also be applied to various decoy-state QKD protocols including MDI-QKD and BB84. Furthermore, we show that a full parameter optimization in MDI-QKD can improve the secure key rate more than one order

of magnitude over non-optimization and it can still increase the key rate around 200% over simplified choice in finite-data settings. Finally, we, for the first time, propose and experimentally implement a novel decoy-state method with only one decoy state. This method is simple to implement, but gives a slightly lower key rate. The protocol and the implementation results are respectively presented in Appendix A 2 and B.

The rest of this paper is organized as follows. We introduce the theory of decoy-state MDI-QKD in Sec. II. In Sec. III, we present our method to perform a full parameter optimization. We show the simulation results about the key rate comparison among full-optimization, partial-optimization and non-optimization in Sec. IV. In Sec. V, we present the simulation results for different number of decoy states. Finally, we conclude this paper in Sec. VI.

II. DECOY-STATE MDI-QKD

The secure key rate of MDI-QKD in the asymptotic case is given by [11]

$$R \geq P_{11}^Z Y_{11}^Z [1 - H_2(e_{11}^X)] - Q_{\mu\mu}^Z f_e(E_{\mu\mu}^Z) H_2(E_{\mu\mu}^Z), \quad (1)$$

where Y_{11}^Z and e_{11}^X are, respectively, the yield (the conditional probability that Charles declares a successful event) in the rectilinear (Z) basis and the error rate in the diagonal (X) basis, given that both Alice and Bob send single-photon states; P_{11}^Z denotes the probability that Alice and Bob send single-photon states in the Z basis; H_2 is the binary entropy function given by $H_2(x) = -x \log_2(x) - (1-x) \log_2(1-x)$; $Q_{\mu\mu}^Z$ and $E_{\mu\mu}^Z$ denote, respectively, the gain and QBER in the Z basis; μ is the intensity of the signal state and its optimal value in the asymptotic case is shown in Appendix D 1; $f_e \geq 1$ is the error correction inefficiency function. Here we use the Z basis for key generation and the X basis for testing only [38]. In practice, $Q_{\mu\mu}^Z$ and $E_{\mu\mu}^Z$ are directly measured in the experiment, while Y_{11}^Z and e_{11}^X can be estimated using the decoy-state methods [27–32].

In a MDI-QKD implementation with coherent states (attenuated lasers), by performing the measurements for different intensity settings, we can obtain [11, 31]

$$Q_{q_a q_b}^\lambda = \sum_{n,m=0}^{\infty} e^{-(q_a+q_b)} \frac{q_a^n}{n!} \frac{q_b^m}{m!} Y_{nm}^\lambda,$$
$$Q_{q_a q_b}^\lambda E_{q_a q_b}^\lambda = \sum_{n,m=0}^{\infty} e^{-(q_a+q_b)} \frac{q_a^n}{n!} \frac{q_b^m}{m!} Y_{nm}^\lambda e_{nm}^\lambda, \quad (2)$$

where $\lambda \in \{X, Z\}$ denotes the basis choice, q_a (q_b) denotes Alice's (Bob's) intensity setting, $Q_{q_a q_b}^\lambda$ ($E_{q_a q_b}^\lambda$) denotes the gain (QBER), and Y_{nm}^λ (e_{nm}^λ) denotes the yield (error rate) given that Alice and Bob send respectively an n-photon and m-photon pulse. Here, the key idea of the finite decoy-state protocol is to estimate a lower bound for Y_{11}^Z and an upper bound for e_{11}^X from the set of linear equations given by Eq. (2). We denote these two bounds as $Y_{11}^{Z,L}$ and $e_{11}^{X,U}$ respectively.

In this work, we focus on the *symmetric* case where the two channel transmissions from Alice to Charles and from Bob to

---

[1] This is because MDI-QKD requires two-fold coincidence events whereas decoy-state BB84 needs only single detection; Given that standard InGaAs single photon detectors in telecom wavelength have a rather low efficiency of say 15%, for a fixed duration of experiment, the data size generated by MDI-QKD is substantially lower than that generated in decoy-state BB84. See the simulation and experimental results [31, 32, 34–37] for more details.

Method	Iterations	Time	Key rate
Exhaustive search	10^7	550 hours	6.84×10^{-5}
Local search	33	1 min	6.83×10^{-5}

TABLE I: Comparison of local search local search algorithm (LSA) and exhaustive search. The simulation is conducted on MDI-QKD with two decoy-state numerical approach (Appendix A 1) using a standard desktop computer. A full optimization on eight dimensions including intensity and probability choices is performed. To reduce the computational complexity of exhaustive search, the intensity of ω is fixed at a near optimal value ω=0.0005 (see Appendix D 3 for the general discussion about the effect of ω). Exhaustive search applies 10 points on seven other dimensions and thus it requires 10^7 iterations for optimization. LSA uses the coordinate descent and backtrack search algorithm [39]. LSA can not only maintain the accuracy of parameter optimization, but can also significantly reduce the computational complexity. The time needed for LSA is four orders of magnitude shorter than an exhaustive search.

Charles are equal. The analysis for asymmetric case can be equivalently conducted by following the techniques presented in [31]. In symmetric case, the optimal intensities for Alice and Bob are *equal* [2]. Hence, to simplify our discussion, we assume that equal intensities are used by Alice and Bob, *i.e.*, $q_a=q_b=q$ with $q \in \{\mu, \nu_1, \nu_2, \omega, ...\}$, where μ denotes the signal state and $\{\nu_1, \nu_2, \omega, ...\}$ denote the decoy states.

This decoy-state estimation can be completed either numerically via linear programming [27, 32] or analytically via gaussian elimination [28–32]. The details of our numerical and analytical methods are respectively presented in Appendix A 1 and A 2.

III. FULL PARAMETER OPTIMIZATION FOR MDI-QKD

In practical QKD applications, for better performance in terms of key rate and distance, it is advantageous to make a serious attempt to optimise operating parameters (*e.g.*, signal/decoy state intensities, basis probabilities and signal/decoy state probabilities). As discussed in Sec. I, the simplified choice (or partial optimization) refers to the basis choice *independent* of the intensity choice. It is commonly used in all previous works on MDI-QKD [27–37]. If Z is used as the majority basis for key generation, the simplified choice will modulate most of signals (over 90%) on Z for all signal and decoy states. Nonetheless, the *key* parameter in a decoy-state estimation is the bit error rate in X, *i.e.*, e_{11}^{X}, which requires a large amount of detection counts for the decoy states in X. The simplified choice, in contrast, results in

a small number of such detection counts and thus increases the estimation error of e_{11}^{X} due to large statistical fluctuations. Therefore, the optimal choice (or full optimization) refers to the basis choice dependent on the intensity choice.

To perform this optimal choice in BB84 and MDI-QKD in the case of two decoy states, we are required to optimize two sets of parameters: intensities of signal and decoy states μ, ν, ω, and the probabilities to choose different intensities and bases $P_\mu, P_\nu, P_{Z|\mu}, P_{Z|\nu}, P_{Z|\omega}$, where P_μ denotes the probability to choose intensity μ and $P_{Z|\mu}$ denotes the conditional probability to choose Z basis conditional on μ. Essentially, it requires a search over eight dimensions [3]. Suppose that a trivial exhaustive search with 10 points on each dimension is conducted, it requires 10^8 iterations to obtain the optimal parameters, which requires over 5000 hours on a standard desktop with 4-core CPUs [4]. At first sight, it might appear to be a hard problem to perform full optimization. However, there is *no* need to perform an exhaustive search.

Here, we introduce a local search algorithm (LSA) [39], a well-known algorithm in the field of computer science, to QKD for this optimization problem. In particular, we adopt the coordinate descent and backtrack search algorithm [39] in our implementation. Coordinate descent can effectively transform a multi-dimensional optimization problem to a one-dimensional line search problem along the direction of one coordinate. This one-dimensional line search problem can be solved by backtrack search algorithm. As a consequence, the LSA enables one to perform a full optimization on all experimental parameters efficiently. We implement this LSA on MDI-QKD and show the comparison results to the trivial exhaustive search in Table I. LSA can be four orders of magnitude faster than a trivial exhaustive search and also achieve a similar optimal key rate. Therefore, LSA cannot only reduce the computational complexity but also present a high accuracy. More details about this algorithm are shown in Appendix C.

η_d	e_d	Y_0	f_e	ϵ	N
14.5%	1.5%	6.02×10^{-6}	1.16	10^{-7}	10^{12}

TABLE II: (Color online) List of practical parameters for numerical simulations. These experimental parameters, including the detection efficiency η_d, the total misalignment error e_d and the background rate Y_0, are from the 144 km QKD experiment reported in [40]. Since two SPDs are used in [40], the background rate of each SPD here is roughly half of the value there. We assume that the four SPDs in MDI-QKD have identical η_d and Y_0. ϵ is the security bound considered in our finite-data analysis. N denotes the total number of signals (weak coherent pulses) sent by Alice and Bob.

---

[2] In this symmetric case, one can prove that Alice's (μ_a) and Bob's (μ_b) optimal signal states in the asymptotic limit satisfy $\mu_a=\mu_b$ by using the model presented in the Appendix B and C1 of [31]; for practical settings, we have performed numerical simulations on all dimensions of parameters (*i.e.*, μ_a, ν_a, ω_a, P_{μ_a},... for Alice and μ_b, ν_b, ω_b, P_{μ_b},... for Bob) and also find that the optimal intensities (and optimal probabilities) are equal.

[3] In the case of Vacuum + weak decoy-state protocol [15, 20, 28, 30], the search can be reduced to six dimensions.

[4] Even though a monte carlo optimization is conducted on a high-performance computer such as the one with 16-core CPUs, it still requires a few days to complete such optimization [22].
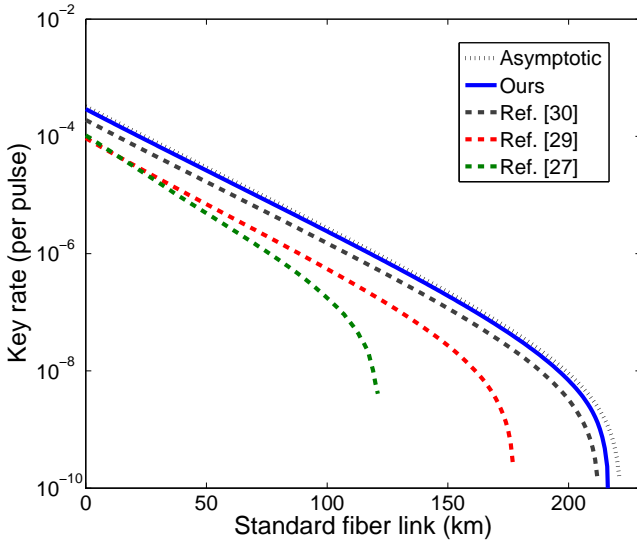
FIG. 1: (Color online) Key rate comparison with infinite data-set. The dotted black curve is the perfect key rate with infinite decoy states. The blue solid curve is our optimized key rate using the numerical approach with two decoy states, where the intensities are $\omega=0.0005$, $\nu=0.01$ and optimized μ. For comparison purpose, we present the non-optimized and partially-optimized key rates using the methods and parameters of Refs [27, 29, 30]: the black dashed curve is using [30] with $\omega=0$, $\nu=0.01$ and optimized μ; the red dashed curve is using [29] with $\omega=0.01$, $\nu=0.1$ and $\mu=0.3$; the green dashed curve is using [27] with $\omega=0$, $\nu=0.1$ and $\mu=0.5$. Notice that if the parameter optimization is also applied to Refs [27, 29], all the key rates are almost the same. In the asymptotic case, parameter optimization is simple, as only the intensities are required to be optimized and a smaller value of decoy-state intensity can in principle result in a better estimation. Parameter optimization can still increase the key rate and extend the secure distance.
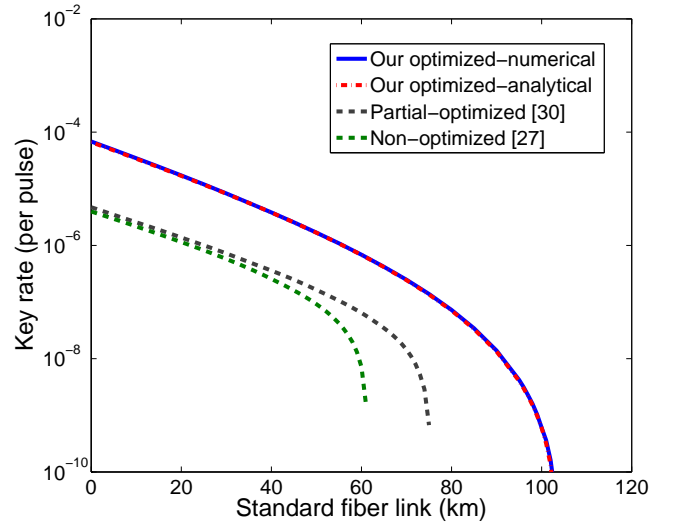

FIG. 2: (Color online) Practical key rate comparison (with statistical fluctuations). The optimal parameters and key rate in the distance of 50km (standard fiber) are shown in Table III. All the key rates are simulated with $N=10^{12}$. The blue solid and red dashed-dotted curves (almost overlapped) are respectively our optimized key rates (after a full optimization) using the numerical (Appendix A 1) and analytical (Appendix A 2) methods with two decoy states. The black dashed curve is using the method of Ref. [30], where only partial parameters (*i.e.*, the intensities) are optimized. The green dashed curve is using the method of Ref. [27], where some typical parameters are assumed without optimization. Without full parameter optimization, the key rates in Refs [27, 30] are around one order of magnitude lower than ours across different distances. Our method can enable secure MDI-QKD over 25km longer than [27, 30]. These results highlight the importance of parameter optimization in practical decoy-state MDI-QKD.

IV. SIMULATIONS ON PARAMETER OPTIMIZATION

In all the simulations presented below, we use the experimental parameters, listed in Table II, mostly from the long-distance QKD experiment reported in [40].

A. Key rate comparison between optimization and non-optimization

For previous works on decoy-state MDI-QKD, Refs. [27, 29] used some typical parameters without optimization and Ref. [30] performed a partial optimization only on intensity choice. Here, we first compare our optimized key rate to those using the parameters and methods presented in Refs. [27, 29, 30]. Fig. 1 shows the comparison results in the asymptotic case. The dotted black curve is the perfect key rate with infinite decoy states. The blue solid curve is the key rate using our numerical method with two decoy states (see Appendix A 1), where we choose the near optimal inten-

sities by maximizing the key rate [5]. The black, red and green dashed curves are respectively using the method and parameters of [30], [29] and [27]. We can see that the key rates without parameter optimization in Refs [27, 29] are much lower than ours and Ref. [30]. Hence, parameter optimization not only increases the key rate but also extends the secure distance in the asymptotic case.

Fig. 2 shows the practical key rates, *i.e.*, with statistical fluctuations, in the case of data-size $N=10^{12}$. The optimal parameters and key rate for the distance of 50km (standard fiber) are shown in Table III. Since Ref. [29] did not consider the finite-data effect, we do not show their key rate here. For a fair comparison, we use the method of standard error analysis [27] to analyze the statistical fluctuations. The key rates without full parameter optimization in Refs [27, 30] are around one or-

---

[5] Notice that in the asymptotic case, the key rate increases with the decrease of the intensity values of decoy states and the probability choice of intensities and basis is not required. To have a fair comparison to [30], we choose the same value of decoy state ν as $\nu=0.01$ and optimize μ. These intensity values can already give a key rate close to the perfect key rate with infinite decoy state.

Parameters	Optimal	Ref. [27]	Ref. [30]
μ	0.25	0.5	0.21
ν	0.05	0.1	0.06
ω	10^{-6}	0	0
P_μ	0.58	0.33	0.33
P_ν	0.30	0.33	0.33
$P_{X|\mu}$	0.03	0.5	0.5
$P_{X|\nu}$	0.71	0.5	0.5
$P_{X|\omega}$	0.83	0.5	0.5
R	1.68×10^{-6}	1.01×10^{-7}	1.64×10^{-7}

TABLE III: Comparison of parameters at 50km standard fiber. More general comparison results are shown in Fig. 2. The 2nd column is the optimal parameters after a full parameter optimization. The 3rd and 4th columns are respectively the parameters from Refs. [27] and [30]. We can see that full optimization can improve the key rate R over one order of magnitude over the non-full-optimization of Refs. [27, 30]. This improvement mainly comes from optimizing the choices of intensities and probabilities. Notice that for the smallest decoy-state ω, modulating the optimal value of around 10^{-6} is usually difficult in decoy-state QKD experiments [16, 17, 19, 21]. However, we find that as long as the intensity of ω is below 1×10^{-3}, the key rate is very close to the optimum (see Appendix D 3 for details).

der of magnitude lower than ours across different distances. Our method can enable secure MDI-QKD over 25km longer than [27, 30]. From Table III, we can see that the improvement mainly comes from the optimization on the choices of intensities and probabilities. We have also performed such comparison at different data-sizes from $N=10^{11}$ to $N=10^{15}$ and the conclusion is almost the same. Note that if the full parameter optimization is also implemented to Refs. [27, 29], all the key rates will be almost the same. These results, once again, highlight the importance of full parameter optimization in the practical implementation of decoy-state MDI-QKD.

B. Key rate comparison between full optimization and partial optimization

Table IV shows the comparison results for different choices of bases. The key rates are simulated using the numerical method with two decoy states. Unbiased denotes the standard protocol with equal basis choice; Simplified denotes the simplified choice with basis choice independent of intensity choice; Optimal denotes the optimal choice with basis choice depending on intensity choice. In a larger data-set of 10^{18} (approaching asymptotic case), the key rates with optimal choice are around 300% higher than those of unbiased choice and close to those of simplified choice. In a reasonable data-set ($N=10^{12}$ to 10^{14}), the key rates with optimal choice are around 300% higher than those of unbiased choice and around 200% higher than those of simplified choice. Therefore, the optimal choice of parameters can significantly increase the key rates in a practical setting with finite data-set.
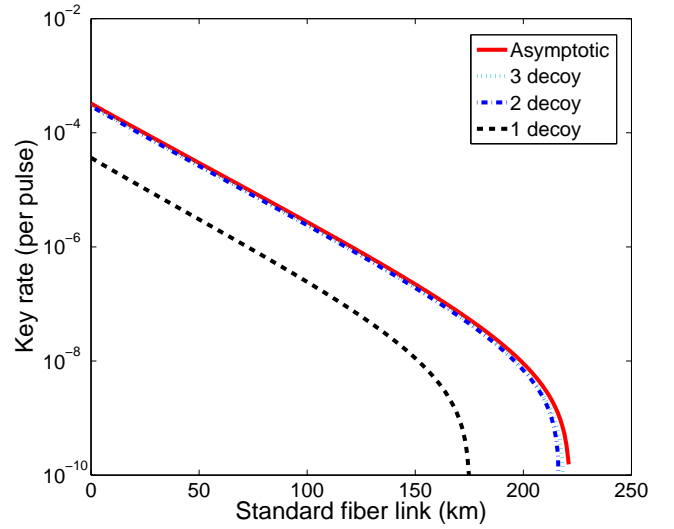

FIG. 3: (Color online) Asymptotic key rates with different number of decoy states. The solid curve is the one with infinite decoy states. The dashed, dashed-dotted, dotted curves are respectively the one, two, three decoy-state results using numerical methods (see Appendix A 1). The signal state μ is optimized in all cases, while some reasonable values of decoy states are adopted: for one decoy state, ν=0.0005; for two decoy states, ν=0.01 and ω=0.0005; for three decoy states, ν_1=0.1, ν_2=0.01 and ω=0.0005. We emphasize that the key rates with analytical methods of Appendix A 2 are almost overlapped with the ones presented in this figure, which shows that the analytical approaches provide a highly good estimation. The estimation using two decoy states gives a nearly similar key rate to the one with three decoy states and is higher than one decoy-state case. Therefore, two decoy states can already result in a near optimal estimation and more decoy states cannot improve the key rate.

V. SIMULATIONS ON DIFFERENT NUMBER OF DECOY STATES

The simulation results using numerical methods (Appendix A 1) for different number of decoy states are shown in Figs. 3, 4 and Table V. In the asymptotic case (Fig. 3), the key rate with two decoy states is close to the one with three decoy states as well as infinite decoy states and it is also larger than that with one decoy state. In a practical setting with finite data-set (Figs. 4), the statistical fluctuations are simulated using the standard error analysis method [15]. A full parameter optimization is conducted using our LSA. Some selected values of key rates are shown in Table V. Our results show that after a full parameter optimization, two decoy states can give an almost optimal key rate, which is much higher than the one with one decoy state. Three decoy states cannot improve the key rates much. Notice that the key rates using the analytical methods of Appendix A 2 are almost overlapped with the ones using numerical methods (see Table V for the case of two decoy states). This shows that the analytical approaches provide a good decoy-state estimation. We have also performed simulations using the rigorous finite-key analysis presented in [32] and find that all the conclusions are the same. Therefore, in practical MDI-QKD, two decoy states combined with full pa-

Distance	0km	0km	0km	50km	50km	50km	100km	100km	100km
Data-size	10^{12}	10^{14}	10^{18}	10^{12}	10^{14}	10^{18}	10^{12}	10^{14}	10^{18}
Unbiased	1.50×10^{-5}	3.98×10^{-5}	6.37×10^{-5}	3.21×10^{-7}	2.39×10^{-6}	5.71×10^{-6}	0	9.88×10^{-8}	4.72×10^{-7}
Simplified	2.05×10^{-5}	6.27×10^{-5}	2.03×10^{-4}	3.36×10^{-7}	3.97×10^{-6}	1.66×10^{-5}	0	1.28×10^{-7}	1.21×10^{-6}
Optimal	6.83×10^{-5}	1.72×10^{-4}	2.72×10^{-4}	1.68×10^{-6}	1.05×10^{-5}	2.24×10^{-5}	6.05×10^{-10}	4.61×10^{-7}	1.78×10^{-6}

TABLE IV: Key rate values with different basis choices. The key rates are simulated with two decoy states and numerical approach. Unbiased denotes the standard protocol with equal basis choice; Simplified denotes the simplified choice with the (biased) basis choice independent of intensity choice; Optimal denotes the optimal choice with the (biased) basis choice depending on intensity choice. In a large data-set of 10^{18} (approaching asymptotic case), the key rates with optimal choice are around 300% higher than those of unbiased choice and close to those of simplified choice. In a reasonable data-set (10^{12} to 10^{14}), the key rates with optimal choice are around 300% higher than those of unbiased choice and around 200% higher than those of simplified choice. This shows that the optimal choice can significantly increase the key rates in a practical setting with finite data-set.

Distance	0km	0km	0km	50km	50km	50km	100km	100km	100km
Data-size	10^{12}	10^{14}	∞	10^{12}	10^{14}	∞	10^{12}	10^{14}	∞
1decoy-Num	8.10×10^{-6}	1.41×10^{-5}	3.59×10^{-5}	2.56×10^{-7}	9.34×10^{-7}	3.03×10^{-6}	0	4.64×10^{-8}	2.41×10^{-7}
2decoy-Num	6.83×10^{-5}	1.72×10^{-4}	3.25×10^{-4}	1.68×10^{-6}	1.05×10^{-5}	2.95×10^{-5}	6.05×10^{-10}	4.61×10^{-7}	2.98×10^{-6}
2decoy-Ana	6.65×10^{-5}	1.67×10^{-4}	$3.25 \times 10{-4}$	1.67×10^{-6}	1.01×10^{-5}	2.95×10^{-5}	5.97×10^{-10}	4.48×10^{-7}	2.77×10^{-6}
3decoy-Num	6.76×10^{-5}	1.71×10^{-4}	3.03×10^{-4}	1.66×10^{-6}	1.04×10^{-5}	2.92×10^{-5}	6.02×10^{-10}	4.55×10^{-7}	2.70×10^{-6}

TABLE V: Optimal key values under different number of decoy states. With finite data-set, the key rates with two decoy states are around one order of magnitude higher than the ones with one decoy state. Three decoy states cannot help to improve the key rates. Hence, two decoy states can achieve a near optimal key rate. In the case of two decoy states, the numerical method (Num) can only improve the key around 2% over the ones using our analytical method (Ana). This shows that the two decoy-state analytical method presented in Appendix A 2 can also result in a near optimal estimation.
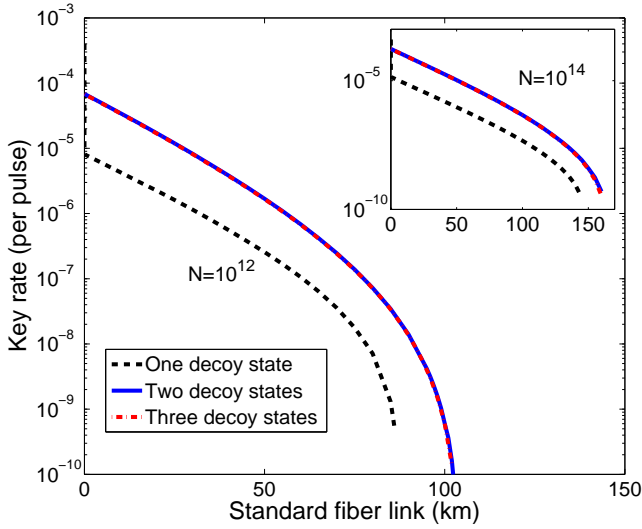

FIG. 4: (Color online) Secret key rate in logarithmic scale as a function of the distance under different numbe of decoy states. The main figure is for data-set $N=10^{12}$ and the inserted figure is for $N=10^{14}$. The key rates are obtained using numerical methods with one (dashed curve), two (solid curve), three (dashed-dotted curve) decoy states. The key rates with two and three decoy states are almost overlapped. In simulation, we perform a full parameter optimization for all cases. Our results show that after a full parameter optimization, the two decoy-state method can give an almost optimal key rate, which is higher than the one with one decoy state. Three decoy states cannot help to increase the key rate.

rameter optimization can achieve a near optimal decoy-state estimation.

VI. CONCLUSION

In summary, we have shown the importance of full parameter optimization in practical decoy-state MDI-QKD and presented a novel LSA to realize such optimization. Full parameter optimization can increase the key rate around 200% over the simplified choice [37]. LSA can be four orders of magnitude faster than a trivial exhaustive search to achieve a similar optimal key rate. In practice, implementing full parameter optimization requires slightly complex modulation schemes, as the sender's two modulators on the intensities and the bit information are dependent. However, one can in principle jointly modulate the two modulators using a single quantum random number generator [41], which is similar to the setup in [34]. Future research can also explore this full parameter optimization in an asymmetric setting of MDI-QKD [31]. Moreover, we have found that two decoy states already enable a near optimal decoy-state estimation in both asymptotic and finite-data case. Experimentalist can readily implement MDI-QKD by using our two decoy-state (analytical or numerical) approach combined with a full parameter optimization to enjoy the optimal system performance.

VII. ACKNOWLEDGMENTS

Appendix A: Finite decoy-state methods

1. Numerical approaches

Ignoring statistical fluctuations temporally, the estimations on $Y_{11}^{Z,L}$ and $e_{11}^{X,U}$ from Eq. (2) are constrained optimisation problems, which is linear and can be efficiently solved by linear programming (LP). The numerical routine to solve these problems can be written as:

$$min : Y_{11}^{Z},$$
$$s.t. : 0 \leq Y_{nm}^{Z} \leq 1, with \, n, m \in \mathcal{S}_{cut}$$
$$Q_{q_a q_b}^{Z} - (1 - \sum_{n,m \in S_{cut}} e^{-(q_a+q_b)} \frac{q_a^n}{n!} \frac{q_b^m}{m!}) \leq$$
$$\sum_{n,m \in \mathcal{S}_{cut}} e^{-(q_a+q_b)} \frac{q_a^n}{n!} \frac{q_b^m}{m!} Y_{nm}^{Z} \leq Q_{q_a q_b}^{Z}$$

$$Max : e_{11}^{X},$$
$$s.t. : 0 \leq Y_{nm}^{X} \leq 1, 0 \leq Y_{nm}^{X} e_{nm}^{X} \leq 1, with \, n, m \in \mathcal{S}_{cut}$$
$$Q_{q_a q_b}^{X} - (1 - \sum_{n,m \in \mathcal{S}_{cut}} e^{-(q_a+q_b)} \frac{q_a^n}{n!} \frac{q_b^m}{m!}) \leq$$
$$\sum_{n,m \in S_{cut}} e^{-(q_a+q_b)} \frac{q_a^n}{n!} \frac{q_b^m}{m!} Y_{nm}^{X} \leq Q_{q_a q_b}^{X}$$
$$Q_{q_a q_b}^{X} E_{q_a q_b}^{X} - (1 - \sum_{n,m \in \mathcal{S}_{cut}} e^{-(q_a+q_b)} \frac{q_a^n}{n!} \frac{q_b^m}{m!}) \leq$$
$$\sum_{n,m \in S_{cut}} e^{-(q_a+q_b)} \frac{q_a^n}{n!} \frac{q_b^m}{m!} Y_{nm}^{X} e_{nm}^{X} \leq Q_{q_a q_b}^{X} E_{q_a q_b}^{X}$$

where $\mathcal{S}_{cut}$ denotes a finite set of indexes n and m, with $\mathcal{S}_{cut} = \{n, m \in \mathbb{N} \text{ with } n \leq N_{cut} and m \leq M_{cut}\}$, for prefixed values of $N_{cut} \geq 2$ and $M_{cut} \geq 2$. In our simulations, we choose $N_{cut} = 7$ and $M_{cut} = 7$, as larger N_{cut} and M_{cut} have negligible effect on decoy-state estimation. More discussions can be seen in [27]. Here, $q \in \{\mu, \nu\}$ for one decoy-state estimation; $q \in \{\mu, \nu, \omega\}$ for two decoy-state estimation; $q \in \{\mu, \nu_1, \nu_2, \omega\}$ for three decoy-state estimation. Notice that statistical fluctuations can be easily conducted by adding constraints on the experimental measurements of $Q_{q_a q_b}^{\lambda}$ and $E_{q_a q_b}^{\lambda}$. These additional constraints can be analyzed by using statistical estimation methods, such as standard error analysis [27] or Chernoff bound [32]. A rigorous finite-key analysis

can also be implemented by following the technique presented in [32].

2. Analytical approaches

a. One decoy state

We consider an estimation method with only one decoy state ν satisfying $\mu > \nu$. Our starting point is Eq. (2). To estimate $Y_{11}^{Z,L}$, we use *gaussian elimination*. Firstly, we simultaneously cancel out all the third order terms Y_{12}, Y_{21}, Y_{30}, Y_{03}:

$$\mu^3 \times Q_{\nu\nu}^{Z} e^{2\nu} - \nu^3 \times Q_{\mu\mu}^{Z} e^{2\mu} =$$
$$\mu^2 \nu^2(\mu-\nu)Y_{11}^{Z} + \mu^3(Y_{00}^{Z} + \nu Y_{01}^{Z} + \nu Y_{10}^{Z} + \nu^2 Y_{02}^{Z}/2 + \nu^2 Y_{20}^{Z}/2)$$
$$- \nu^3(Y_{00}^{Z} + \mu Y_{01}^{Z} + \mu Y_{10}^{Z} + \mu^2 Y_{02}^{Z}/2 + \mu^2 Y_{20}^{Z}/2) +$$
$$\sum_{n+m>3}^{\infty} \frac{(\nu^{n+m}\mu^3 - \mu^{n+m}\nu^3)}{n!m!} Y_{nm}^{Z} \leq$$
$$\mu^2 \nu^2(\mu-\nu)Y_{11}^{Z} + \mu^3(Y_{00}^{Z} + \nu Y_{01}^{Z} + \nu Y_{10}^{Z} + \nu^2 Y_{02}^{Z}/2 + \nu^2 Y_{20}^{Z}/2)$$
$$\tag{A1}$$

where the inequality comes from the fact that $(\nu^{n+m}\mu^3 - \mu^{n+m}\nu^3) < 0$ for $n + m > 3$. Next, from $Q_{\nu\nu}^{Z} E_{\nu\nu}^{Z}$, we have

$$Q_{\nu\nu}^{Z} E_{\nu\nu}^{Z} e^{2\nu} = \sum_{n,m=0}^{\infty} \frac{\nu^{n+m}}{n!m!} Y_{nm}^{Z} e_{nm}^{Z} \geq$$
$$Y_{00}^{Z} e_{00}^{Z} + \nu Y_{01}^{Z} e_{01}^{Z} + \nu Y_{10}^{Z} e_{10}^{Z} + \nu^2 Y_{02}^{Z} e_{02}^{Z}/2 + \nu^2 Y_{20}^{Z} e_{20}^{Z}/2$$
$$= (Y_{00}^{Z} + \nu Y_{01}^{Z} + \nu Y_{10}^{Z} + \nu^2 Y_{02}^{Z}/2 + \nu^2 Y_{20}^{Z}/2)/2 \tag{A2}$$

where the final equality is from $e_{0m}^{Z} = e_{n0}^{Z} = 1/2$, which is a standard assumption in QKD descending from the fact that the error rate cause by 0-photon pulse is 1/2. Therefore, by combining Eq. (A1) and Eq. (A2), we have a lower bound for Y_{11}^{Z}

$$Y_{11}^{Z} \geq Y_{11}^{Z,L} = \frac{\mu^3 Q_{\nu\nu}^{Z} e^{2\nu}(1 - 2E_{\nu\nu}^{Z}) - \nu^3 Q_{\mu\mu}^{Z} e^{2\mu}}{\mu^2 \nu^2(\mu - \nu)} \tag{A3}$$

To estimate $e_{11}^{X,U}$, we use the same method as [31, 32] and obtain an upper bound for e_{11}^{X}

$$e_{11}^{X} \leq e_{11}^{X,U} = \frac{1}{(\mu-\nu)^2 Y_{11}^{X,L}} \times$$
$$(e^{2\mu} Q_{\mu\mu}^{X} E_{\mu\mu}^{X} + e^{2\nu} Q_{\nu\nu}^{X} E_{\nu\nu}^{X} - e^{\mu+\nu} Q_{\mu\nu}^{X} E_{\mu\nu}^{X} - e^{\nu+\mu} Q_{\nu\mu}^{X} E_{\nu\mu}^{X}) \tag{A4}$$

b. Two decoy states

We consider an estimation method with two decoy states ν, ω satisfying $\mu > \nu > \omega \geq 0$. We have the lower bound $Y_{11}^{Z,L}$

and the upper bound $e_{11}^{X,U}$ [31, 32]

$$Y_{11}^{Z,L} = \frac{1}{(\mu-\omega)^2(\nu-\omega)^2(\mu-\nu)} \times$$
$$[(\mu^2-\omega^2)(\mu-\omega)(Q_{\nu\nu}^Z e^{2\nu}+Q_{\omega\omega}^Z e^{2\omega}-Q_{\nu\omega}^Z e^{\nu+\omega}-Q_{\omega\nu}^Z e^{\omega+\nu})$$
$$-(\nu^2-\omega^2)(\nu-\omega)(Q_{\mu\mu}^Z e^{2\mu}+Q_{\omega\omega}^Z e^{2\omega}-Q_{\mu\omega}^Z e^{\mu+\omega}-Q_{\omega\mu}^Z e^{\omega+\mu})],$$
$$(A5)$$

$$e_{11}^{X,U} = \frac{1}{(\nu-\omega)^2 Y_{11}^{X,L}} \times$$
$$[e^{2\nu}Q_{\nu\nu}^X E_{\nu\nu}^X+e^{2\omega}Q_{\omega\omega}^X E_{\omega\omega}^X-e^{\nu+\omega}Q_{\nu\omega}^X E_{\nu\omega}^X-e^{\omega+\nu}Q_{\omega\nu}^X E_{\omega\nu}^X].$$
$$(A6)$$

Appendix B: Experimental implementation of the one decoy-state method

Here we experimentally implement the one decoy-state method (Appendix A 2) in a polarization-encoding MDI-QKD system presented in [37]. Using the systematic parameters shown in Table VI, we perform a numerical optimization to maximize the key rate. The optimal intensities of the signal and the decoy state are respectively around $\mu = 0.1$ and $\nu = 0.01$ and the optimal probability to send a signal state is around $P_s = 0.45$. Here we choose the simplified choice with the same probability to select Z or X basis, which is simpler for our implementation. An optimal choice will be implemented in our future experiments.

We test the one decoy-state method over 10km standard telecom fiber and operate the system at a repetition rate around 500 kHz. The system is operated for 55 hours and a total number of signals around $N = 1.11 \times 10^{11}$ is generated. The experimental results are shown in Table VI. Around 50 secure keys are generated. Our results demonstrate the possibility of one decoy-state method. This method is highly simple to implement in a practical MDI-QKD system, but gives a slightly lower key rate.

One might ask 'why the key rate is low in our implementation?'. This is mainly due to the low repetition rate of our system as well as the finite-key effect. The key generation rate can be substantially improved by increasing the repetition rate: First, more pulses can be sent out in a reasonable time frame, leading to tighter bounds in the decoy-state estimation. Second, by using four detectors rather than two (as our implementation), we can get at least a four-fold increase in the key rate. Third, given a larger data size, we can reduce the portion of pulses sent as decoy states and more pulses can be sent out in signal states for key generation. The speed of our system is limited by the performance of our SPDs. Our simulation (see the black dashed curves for the asymptotic case in Fig. 3 and the finite-key case with $N = 10^{14}$ in Fig. 4) shows that, with commercial four single-photon detectors having say
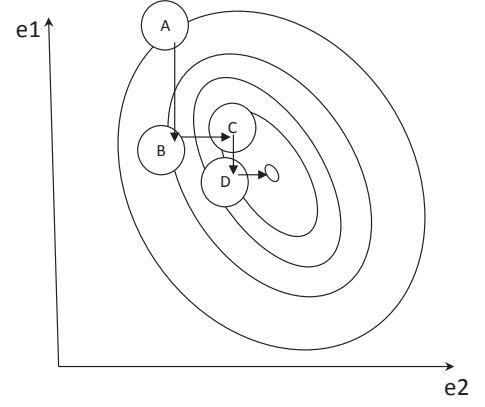

FIG. 5: Coordinate Descent (CD). CD algorithm searches along one coordinate direction in each iteration, and it use a different coordinate directions cyclically. For instance, on the equiv-error contour of two dimensional subspace, CD starts at point A (arbitrarily) and descents vertically along the direction e_1 to B, then horizontally along the direction e_2 toward C. After cyclic iterations of vertical and horizontal descent, the algorithm stops at D where it is very close to the optimal. This simplified two-dimensional example illustrates how generalized search in any dimensional space can be done analogously.

an efficiency of about 15% [6], the key generation rate can be around 10^{-5} per pulse at say 30km fiber. If gating up to 100 MHz, the key rate (with the finite key effect) can be up to 1 kbps. Moreover, by using state-of-the-art super-conducting single photon detectors with over 90% quantum efficiency, the key rate can be as high as 100 kbps. This high-speed system will be implemented in our future experiment.

Appendix C: Local search algorithm

In the decoy-state approach to BB84 and MDI-QKD, the key rate depends largely on the systematic parameters and the optimal key rate is achieved via numerical optimization on many dimensions (parameters). To reduce both computational time and storage space of this optimization process, local search algorithm (LSA), a combination of coordinate descent (CD) and backtrack search (BS) algorithm, is adopted in lieu of the conventional exhaustive search algorithm.

There are salient drawbacks of exhaustive search. If the search is too fine, the computational time and space are challenging. If the search is too coarse, we will miss finer details. In contrast, CD is a non-derivative approximation to the well-known steepest descent (SD) algorithm [42]. This approximation is necessary from the facts that our key rate is an implicit function (a linear program) of the parameters and the actual finding of gradients and hessians of SD cannot be done easily. CD converges to the same optimal point as SD, even though

---

[6] For instance, id210/220, manufactured by IDQ, can have over 20% efficiency and a gate rate over 100 MHz.

Systematic parameters	η_d	e_d	Y_0	L	ϵ	N
	8.2%	0.8%	5×10^{-5}	10km	2.7×10^{-3}	1.11×10^{11}
One decoy state	P_s	$Q_{\mu\mu}^{Z}$	$E_{\mu\mu}^{Z}$	$Y_{11}^{Z,L}$	$e_{11}^{X,U}$	R_L
	0.45	5.33×10^{-6}	4.03%	4.17×10^{-5}	12.97%	4.26×10^{-10}

TABLE VI: Experimental parameters and results. These experimental parameters include the detection efficiency η_d, the total misalignment error e_d, the background rate Y_0, the transmission distance from Alice to Bob L, the security bound ϵ and the total number of signals sent by Alice and Bob. P_s is the optimal probability to send a signal state. R_L denotes the lower bound of the key rate.

it requires more iterations. CD can fix low-speed by making large progress at the start, and it can also fix in-accuracy by re-defining how close to the optimal point the algorithm can stop. Table I compares the speed and accuracy of exhaustive search versus LSA.

CD is based on the idea that the minimization of a multi-variate function (key rate R) can be achieved by minimizing it along one direction at a time (see Fig. 5). Instead of varying descent direction according to gradient, one fixes descent directions at the outset [43]. These directions are usually the cartesian bases, *i.e.*, e_i with i=1,2,3,.... In the two decoy-state case, $e_1=\mu$, ..., $e_4=P_\mu$, ..., $e_7=P_{X|\mu}$, $e_8=P_{X|\nu}$, and this basis is iterated through one at a time, minimizing the objective function with respect to the current coordinate direction. Mathematically, to optimize μ, if μ^k (optimized μ in the kth iteration) is given, the minimization of key rate R (see Eq. 1) along μ coordinate in the $k + 1$th iteration is:

$$\mu^{k+1} = arg \max_{y \in R} R(P_\mu^{k+1}, P_\nu^{k+1},$$
$$P_{X|\mu}^{k+1}, P_{X|\nu}^{k+1}, P_{X|\omega}^{k+1}, y, \nu^k, \omega^k) \quad \text{(C1)}$$

By doing line search in each iteration, we automatically have a sequence of vector $x_0, x_1, x_2, ...$, where $x_i = ((P_\mu)_i, (P_\nu)_i, ..., (P_{X|\mu})_i, ..., (\nu)_i, (\omega)_i)$ and the sequence of key rate: $R(x_0) \geq R(x_1) \geq R(x_2) \geq ...$.

The demonstration of convexity in Appendix D 2 will make the result of LSA a global optimum. Although CD requires an intelligent guess to start with, the starting point in convex topologies can be in theory any non-zero objective (key rate) point in the search space. In practice, prior research can shed light on the choices of initial parameters, and these parameters often are good candidates for the starting guess.

After a direction along a coordinate chosen in CD, we still have to do a one-dimensional line search problem to compute how far the search can move along a given coordinate. This is realized via the BS algorithm. BS starts at the end of previous iterations, and makes progress toward a minima along the chosen coordinate direction. With a step from one side to the minima to the other side, the algorithm found a turning point. From there, BS searches backward again toward the minima until the same turning point is found with greater accuracy. The procedure is iterated until converged to the minima. At this point in the search space, the CD algorithm restarts with a new direction of line-search.

Appendix D: Other practical aspects

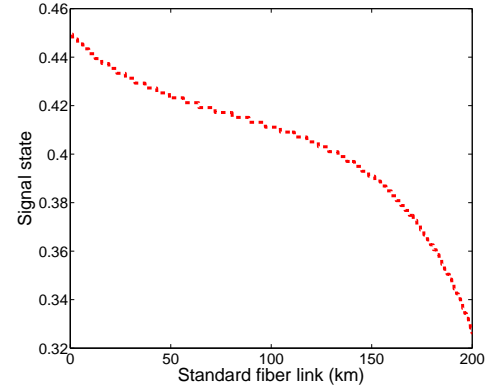**1. Optimal signal state in the asymptotic case**

FIG. 6: Plot of the intensity μ as a function of the transmission distance for the decoy-state MDI-QKD with infinite decoy states.

We consider the asymptotic case with infinite decoy states. From the model presented in [31], we have:

$$e_{11}^{X} \approx e_d$$
$$Y_{11}^{Z} \approx t_a t_b \eta_d^2$$
$$Q_{\mu\mu}^{Z} \approx \mu^2 t_a t_b \eta_d^2 (1 + 2e_d)/2 \quad \text{(D1)}$$
$$E_{\mu\mu}^{Z} \approx \frac{2e_d}{1 + e_d}$$

where t_a (t_b) is the channel transmittance from Alice (Bob) to Charlie.

Substituting Eq. (D1) into Eq. (1), the key generation rate is given by:

$$R \geq \mu^2 e^{-2\mu} t_a t_b \eta_d^2 [1 - H_2(e_d)] -$$
$$\frac{\mu^2 t_a t_b \eta_d^2 (1 + 2e_d)}{2} f_e(\frac{2e_d}{1 + e_d}) H_2(\frac{2e_d}{1 + e_d}), \quad \text{(D2)}$$

The expression is optimized if we choose $\mu = \mu^{Optimal}$ which fulfills $\frac{\partial R}{\partial \mu} = 0$:

$$(1 - \mu) \exp(-2\mu) = \frac{f_e(\frac{2e_d}{1+e_d})(1 + 2ed) H_2(\frac{2e_d}{1+e_d})}{1 - H_2(e_d)}. \quad \text{(D3)}$$

By using $f_e(\cdot) = 1.16$ and $e_d = 1.5\%$, we solve Eq. (D3) and get $\mu^{Optimal} = 0.42$. The numerical simulation for the

optimal μ at different distances is shown in Fig. 6. We can see that Eq. (D3) is a good approximation. Moreover, from Eq. (D3) and Fig. 6, the optimal intensity is a continuous function with transmission distance and there is *only* one solution for $\mu \in (0, 1]$. That is, the key rate is a *convex function* to μ at a fixed distance.

2. Convex function of the key rate

The CD and BS algorithm [39] works in any general topology of the search space, but the saving in efficiency comes only when the underlying topology is a *convex optimization* problem. It is our goal here to demonstrate the convexity of the key rate as a function of the parameters. Appendix D 1 has already shown that the key rate is indeed convex in the case of infinite decoy states. In a practical setting with finite decoy states, for illustration purpose, we have chosen to sweep the intensities of μ and ν and optimized the other dimensions at

0km with $N=10^{12}$. Fig. 7(a) shows that the convexity of key rate function, which allows a unique optimal set of parameters to be employed in an actual experiment.

3. The effect of the smallest decoy state ω

In practice, it is usually difficult to create a perfect vacuum state in decoy-state QKD experiments [17, 19]. The different intensities are usually generated with an intensity modulator, which has a finite extinction ratio below 30 dB. Thus, the question is: what is the effect of the intensity of the smallest decoy states ω on the secure key rate? Here, we perform a simulation on the sweep of the smallest intensity ω in the case of two decoy states. The result is shown in Fig. 7(b). We find that after a full optimization of parameters, the optimal ω is in the vicinity of 5×10^{-6}. As long as the intensity of ω is below 1×10^{-3}, the key rate is very close to the optimum. In summary, a perfect vaccum ($\omega = 0$) is *not* essentially required in practical decoy-state QKD experiments.

* Electronic address: feihu.xu@utoronto.ca

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

[2] A. K. Ekert, Physical Review Letters **67**, 661 (1991).

[3] M. Peev, C. Pacher, R. Alléaume, C. Barreiro, J. Bouda, W. Boxleitner, T. Debuisschert, E. Diamanti, M. Dianati, J. Dynes, *et al.*, New Journal of Physics **11**, 075001 (2009).

[4] M. Sasaki, M. Fujiwara, H. Ishizuka, W. Klaus, K. Wakui, M. Takeoka, S. Miki, T. Yamashita, Z. Wang, A. Tanaka, *et al.*, Optics Express **19**, 10387 (2011).

[5] Y. Zhao, C. Fung, B. Qi, C. Chen, and H.-K. Lo, Physical Review A **78**, 042333 (2008).

[6] F. Xu, B. Qi, and H.-K. Lo, New Journal of Physics **12**, 113026 (2010).

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

[8] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Nature Communications **2**, 349 (2011).

[9] H. Weier, H. Krauss, M. Rau, M. Fuerst, S. Nauerth, and H. Weinfurter, New Journal of Physics **13**, 073024 (2011).

[10] N. Jain, C. Wittmann, L. Lydersen, C. Wiechers, D. Elser, C. Marquardt, V. Makarov, and G. Leuchs, Physical Review Letters **107**, 110501 (2011).

[11] H.-K. Lo, M. Curty, and B. Qi, Physical Review Letters **108**, 130503 (2012).

[12] W. Hwang, Physical Review Letters **91**, 57901 (2003).

[13] H.-K. Lo, X. Ma, and K. Chen, Physical Review Letters **94**, 230504 (2005).

[14] X. Wang, Physical Review Letters **94**, 230503 (2005).

[15] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, Physical Review A **72**, 012326 (2005).

[16] Y. Zhao, B. Qi, X. Ma, H.-K. Lo, and L. Qian, Physical review letters **96**, 070502 (2006).

[17] D. Rosenberg, J. W. Harrington, P. R. Rice, P. A. Hiskett, C. G. Peterson, R. J. Hughes, A. E. Lita, S. W. Nam, and J. E. Nordholt, Physical Review Letters **98**, 010503 (2007).

[18] C.-Z. Peng, J. Zhang, D. Yang, W.-B. Gao, H.-X. Ma, H. Yin,

H.-P. Zeng, T. Yang, X.-B. Wang, and J.-W. Pan, Physical review letters **98**, 010505 (2007).

[19] A. Dixon, Z. Yuan, J. Dynes, A. Sharpe, and A. Shields, Optics Express **16**, 18790 (2008).

[20] Z. Wei, W. Wang, Z. Zhang, M. Gao, Z. Ma, and X. Ma, Scientific reports **3**, 2453 (2013).

[21] M. Lucamarini, K. Patel, J. Dynes, B. Fröhlich, A. Sharpe, A. Dixon, Z. Yuan, R. Penty, and A. Shields, Optics express **21**, 24550 (2013).

[22] C. C. W. Lim, M. Curty, N. Walenta, F. Xu, and H. Zbinden, Phys. Rev. A **89**, 022307 (2014).

[23] Z. Yuan, J. Dynes, and A. Shields, Applied physics letters **98**, 231104 (2011).

[24] D. Mayers and A. Yao, Quantum Information & Computation **4**, 273 (2004).

[25] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Physical Review Letters **98**, 230501 (2007).

[26] S. L. Braunstein and S. Pirandola, Physical Review Letters **108**, 130502 (2012).

[27] X. Ma, C.-H. F. Fung, and M. Razavi, Physical Review A **86**, 052305 (2012).

[28] X.-B. Wang, Physical Review A **87**, 012320 (2013).

[29] Z.-W. Yu, Y.-H. Zhou, and X.-B. Wang, arXiv preprint arXiv:1309.5886 (2013).

[30] S.-H. Sun, M. Gao, C.-Y. Li, and L.-M. Liang, Physical Review A **87**, 052329 (2013).

[31] F. Xu, M. Curty, B. Qi, and H.-K. Lo, New Journal of Physics **15**, 113007 (2013).

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

[33] F. Xu, B. Qi, Z. Liao, and H.-K. Lo, Applied Physics Letters **103**, 061101 (2013).

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

[35] Y. Liu, T.-Y. Chen, L.-J. Wang, H. Liang, G.-L. Shentu, J. Wang, K. Cui, H.-L. Yin, N.-L. Liu, L. Li, X. Ma, J. S. Pelc, M. M. Fejer, C.-Z. Peng, Q. Zhang, and J.-W. Pan, Phys. Rev. Lett. **111**, 130502 (2013).
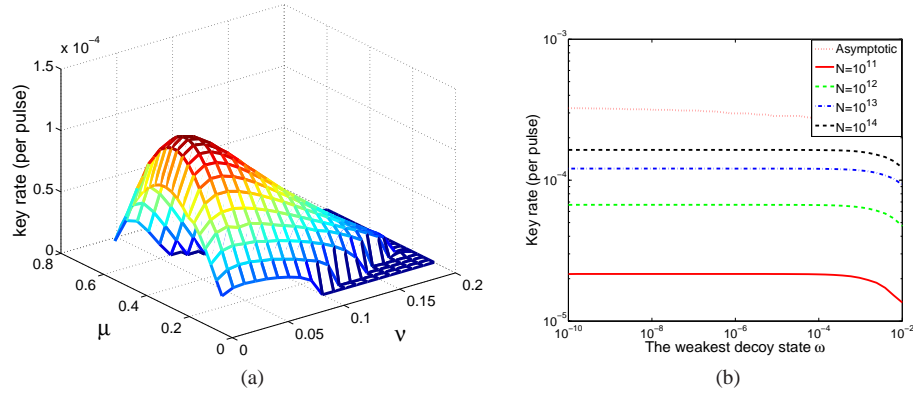
FIG. 7: (Color online) (a) Convexity of key rate function. The key rate is simulated by sweeping μ and ν, and optimizing other parameters. (b) Key rate as a function of the decoy state ω. As long as the intensity of ω is below 1×10^{-3}, the key rate is very close to the optimum. A perfect vaccum ($\omega = 0$) is not essentially required in practical decoy-state QKD experiments.

[36] T. Ferreira da Silva, D. Vitoreti, G. B. Xavier, G. C. do Amaral, G. P. Temporão, and J. P. von der Weid, Phys. Rev. A **88**, 052303 (2013).

[37] Z. Tang, Z. Liao, F. Xu, B. Qi, L. Qian, and H.-K. Lo, arxiv:1306.6104 [to be published in Physics Review Letters] (2014).

[38] H.-K. Lo, H.-F. Chau, and M. Ardehali, Journal of Cryptology **18**, 133 (2005).

[39] S. P. Boyd and L. Vandenberghe, *Convex optimization* (Cambridge university press, 2004).

[40] R. Ursin, F. Tiefenbacher, T. Schmitt-Manderbach, H. Weier, T. Scheidl, M. Lindenthal, B. Blauensteiner, T. Jennewein, J. Perdigues, P. Trojek, B. Ömer, *et al.*, Nature Physics **3**, 481 (2007).

[41] F. Xu, B. Qi, X. Ma, H. Xu, H. Zheng, and H.-K. Lo, Optics Express **20**, 12366 (2012).

[42] Z.-Q. Luo and P. Tseng, Journal of Optimization Theory and Applications **72**, 7 (1992).

[43] D. P. Bertsekas, Athena Scientific (1999).