Practical decoy state for quantum key distribution

Xiongfeng Ma, Bing Qi, Yi Zhao, and Hoi-Kwong Lo
Center for Quantum Information and Quantum Control, Department of Physics and Department of Electrical & Computer Engineering, University of Toronto, Toronto, Ontario, Canada, M5S 3G4
(Received 7 March 2005; published 20 July 2005)

Decoy states have recently been proposed as a useful method for substantially improving the performance of quantum key distribution (QKD). Here, we present a general theory of the decoy state protocol based on only two decoy states and one signal state. We perform optimization on the choice of intensities of the two decoy states and the signal state. Our result shows that a decoy state protocol with only two types of decoy states—the vacuum and a weak decoy state—asymptotically approaches the theoretical limit of the most general type of decoy state protocol (with an infinite number of decoy states). We also present a one-decoy-state protocol. Moreover, we provide estimations on the effects of statistical fluctuations and suggest that, even for long-distance (larger than 100 km) QKD, our two-decoy-state protocol can be implemented with only a few hours of experimental data. In conclusion, decoy state quantum key distribution is highly practical.

DOI: 10.1103/PhysRevA.72.012326 PACS number(s): 03.67.Dd

I. INTRODUCTION

The goal of quantum key distribution (QKD) [1] is to allow two distant parties Alice and Bob to share a common string of secret data (known as the key), in the presence of an eavesdropper Eve. Unlike conventional cryptography, QKD promises perfect security based on the fundamental laws of physics. Proving the unconditional security of QKD is a hard problem. Fortunately, this problem has recently been solved [2,3]. See also [4]. Experimental QKD has been successfully demonstrated over 100 km of commercial Telecom fibers [5,6] and commercial QKD systems are already on the market. The most important question of QKD is its security. Real-life QKD systems are often based on attenuated laser pulses (i.e., weak coherent states), which occasionally give out more than one photon. This opens up the possibility of sophisticated eavesdropping attacks such as a photon number splitting attack, where Eve stops all single-photon signals and splits multiphoton signals, keeping one copy herself and resending the rest to Bob. The security of practical QKD systems has previously been discussed in [7].

Hwang [8] proposed the decoy state method as an important weapon to combat such a sophisticated attack: by preparing and testing the transmission properties of some decoy states, Alice and Bob are in a much better position to catch an eavesdropper. Hwang specifically proposed to use a decoy state with an average number of photons of order 1. Hwang's idea was highly innovative. However, his security analysis was heuristic.

In [9], we presented a rigorous security analysis of the decoy state idea. More specifically, we combined the idea of the entanglement distillation approach by Gottesman, Lo, Lutkenhaus, and Preskill (GLLP) [7] with the decoy method and achieved a formula for key generation rate:

$$R \geqslant q\{-Q_\mu f(E_\mu)H_2(E_\mu) + Q_1[1 - H_2(e_1)]\}, \qquad (1)$$

where q depends on the implementation [1/2 for the Bennett-Brassard 1984 (BB84) protocol due to the fact that half of the time Alice and Bob disagree with the bases, and if one uses the efficient BB84 protocol [10], $q \approx 1$], the subscript μ denotes the intensity of signal states, Q_μ is the gain [11] of signal states, E_μ is the overall quantum bit error rate (QBER), Q_1 is the gain of single-photon states, e_1 is the error rate of single-photon states, $f(x)$ is the bidirectional error correction efficiency (see, for example, [12]) as a function of error rate, normally $f(x) \geqslant 1$ with Shannon limit $f(x)=1$, and $H_2(x)$ is the binary Shannon information function, given by

$$H_2(x) = -x \log_2(x) - (1-x)\log_2(1-x).$$

Four key variables are needed in Eq. (1). Q_μ and E_μ can be measured directly from the experiment. Therefore, in the paper [9], we showed rigorously how one can, by using the decoy state idea to estimate Q_1 and e_1, thus achieve the unconditional security of QKD with the key generation rate given by Eq. (1). Moreover, using the experimental parameters from a particular QKD experiment [that of Gobby, Yuan, and Shields (GYS)] [5], we showed that decoy state QKD can be secure over 140 km of Telecom fibers. In summary, we showed clearly that the decoy state can indeed substantially increase both the distance and the key generation rate of QKD.

For practical implementations, we also emphasized that only a few decoy states will be sufficient [9]. This is so because contributions from states with large photon numbers are negligible in comparison with those from small photon numbers. In particular, we proposed a vacuum+weak decoy state protocol. That is to say, there are two decoy states—a vacuum and a weak decoy state. Moreover, the signal state is chosen to be of order 1 photon on average. The vacuum state is particularly useful for estimating the background detection rate. Intuitively, a weak decoy state allows us to lower-bound Q_1 and upper-bound e_1.

Subsequently, the security of our vacuum+weak decoy state protocol has been analyzed by Wang [13]. Let us denote the intensities of the signal state and the nontrivial decoy state by μ and μ', respectively. Wang derived a useful upper bound for Δ:

$$\Delta \leqslant \frac{\mu}{\mu' - \mu}\left(\frac{\mu e^{-\mu} Q_{\mu'}}{\mu' e^{-\mu'} Q_{\mu}} - 1\right) + \frac{\mu e^{-\mu} Y_0}{\mu' Q_{\mu}}, \tag{2}$$

where Δ is the proportion of "tagged" states in the sifted key as defined by GLLP [7]. Whereas we [9] considered a strong version of the GLLP result noted in Eq. (1), Wang proposed to use a weak version of the GLLP result:

$$R \geqslant q Q_{\mu}\left\{-H_2(E_{\mu}) + (1-\Delta)\left[1 - H_2\left(\frac{E_{\mu}}{1-\Delta}\right)\right]\right\}. \tag{3}$$

Such a weak version of the GLLP result does not require an estimation of e_1. So it has the advantage that the estimation process is simple. However, it leads to lower values of the key generation rates and distances. The issue of statistical fluctuations in decoy state QKD was also mentioned in [13].

Our observation [9] that only a few decoy states are sufficient for practical implementations has been studied further and confirmed in a recent paper [14].

The main goal of this paper is to analyze the security of a rather general class of two-decoy-state protocols with two weak decoy states and one signal state. Our main contributions are as follows. First, we derive a general theory for a decoy state protocol with two weak decoy states. Whereas Wang [13] considered only our vacuum+weak decoy state protocol [9] (i.e., a protocol with two decoy states—the vacuum and a weak coherent state), our analysis here is more general. Our decoy method applies even when both decoy states are nonvacuum. Note that, in practice, it may be difficult to prepare a vacuum decoy state. For instance, standard variable optical attenuators (VOAs) cannot block optical signals completely. For the special case of the vacuum+weak decoy state protocol, our result generalizes the work of Wang [13].

Second, we perform an optimization of the key generation rate in Eq. (1) as a function of the intensities of the two decoy states and the signal state. Up till now, such an optimization problem has been a key unresolved problem in the subject. We solve this problem analytically by showing that the key generation rate given by Eq. (1) is optimized when both decoy states are weak. In fact, in the limit that both decoy states are infinitesimally weak, we match the best lower bound on Y_1 and upper bound on e_1 in the most general decoy state theory where an infinite number of decoy states are used. Therefore, asymptotically, there is no obvious advantage in using more than two decoy states.

Third, for practical applications, we study the correction terms to the key generation rate when the intensities of the two decoy states are nonzero. We see that the correction terms (to the asymptotically zero-intensity case) are reasonably small. For the case where one of the two decoy states is a vacuum (i.e., $v_2=0$), the correction term remains modest even when the intensity of the second decoy state, ν_1, is as high as 25% of that of the signal state.

Fourth, following [13], we discuss the issue of statistical fluctuations due to finite data size in real-life experiments. We provide a rough estimation on the effects of statistical fluctuations in practical implementations. Using a recent experiment [5] as an example, we estimate that our weak decoy state proposal with two decoy states (a vacuum and a weak decoy state of strength ν) can achieve secure QKD over more than 100 km with only a few hours of experiments. A caveat of our investigation is that we have not considered the fluctuations in the intensities of Alice's laser pulses (i.e., the values of μ, ν_1, and ν_2). This is mainly because of a lack of reliable experimental data. In summary, our result demonstrates that our two-decoy-state proposal is highly practical.

Fifth, we also present a one-decoy-state protocol. Such a protocol has the advantage of being simple to implement, but gives a lower key generation rate. Indeed, we have recently demonstrated experimentally our one-decoy-state protocol over 15 km [15]. This demonstrates that one decoy state is, in fact, sufficient for many practical applications. In summary, decoy state QKD is simple and cheap to implement and it is, therefore, ready for immediate commercialization.

We remark in passing that a different approach (based on a strong reference pulse) to making another protocol (B92 protocol) [16]unconditionally secure over a long distance has recently been proposed in a theoretical paper by Koashi [17].

The organization of this paper is as follows. In Sec. II, we model an optical-fiber-based QKD setup. In Sec. III, we first give a general theory for m decoy states. We then propose our practical decoy method with two decoy states. In addition, we optimize our choice of the average photon numbers μ of the signal state and ν_1 and ν_2 of the decoy states by maximizing the key generation rate with the experimental parameters in a specific QKD experiment (GYS) [5]. Furthermore, we also present a simple one-decoy-state protocol. In Sec. IV, we discuss the effects of statistical fluctuations in the two-decoy-state method for finite data-set size (i.e., the number of pulses transmitted by Alice). Finally, in Sec. V, we present some concluding remarks.

II. MODEL

In order to describe a real-world QKD system, we need to model the source, channel, and detector. Here we consider a widely used fiber-based setup model [18].

A. Source

The laser source can be modeled as a weak coherent state. Assuming that the phase of each pulse is totally randomized, the photon number of each pulse follows a Poisson distribution with a parameter μ as its expected photon number set by Alice. Thus, the density matrix of the state emitted by Alice is given by

$$\rho_A = \sum_{i=0}^{\infty} \frac{\mu^i}{i!} e^{-\mu} |i\rangle\langle i|, \tag{4}$$

where $|0\rangle\langle 0|$ is the vacuum state and $|i\rangle\langle i|$ is the density matrix of the i-photon state for $i=1,2,\ldots$.

B. Channel

For an optical-fiber-based QKD system, the losses in the quantum channel can be derived from the loss coefficient α measured in dB/km and the length of the fiber l in km. The

channel transmittance t_{AB} can be expressed as

$$t_{AB} = 10^{-\alpha l/10}.$$

C. Detector

Let η_{Bob} denote the transmittance in Bob's side, including the internal transmittance of optical components t_{Bob} and detector efficiency η_D,

$$\eta_{Bob} = t_{Bob}\eta_D.$$

Then the overall transmission and detection efficiency between Alice and Bob η is given by

$$\eta = t_{AB}\eta_{Bob}. \tag{5}$$

It is common to consider a *threshold* detector in Bob's side. That is to say, we assume that Bob's detector can tell a vacuum from a nonvacuum state. However, it cannot tell the actual photon number in the received signal, if it contains at least one photon.

It is reasonable to assume independence between the behaviors of the i photons in an i-photon state. Therefore the transmittance of the i-photon state η_i with respect to a threshold detector is given by

$$\eta_i = 1 - (1 - \eta)^i \tag{6}$$

for $i=0,1,2,\ldots$.

D. Yield

Define Y_i to be the yield of an i-photon state, i.e., the conditional probability of a detection event at Bob's side given that Alice sends out an i-photon state. Note that Y_0 is the background rate, which includes the detector dark count and other background contributions such as the stray light from timing pulses.

The yield of an i-photon state, Y_i, mainly comes from two parts, background and true signal. Assuming that the background counts are independent of the signal photon detection, then Y_i is given by

$$Y_i = Y_0 + \eta_i - Y_0\eta_i \cong Y_0 + \eta_i. \tag{7}$$

Here we assume that Y_0 (typically 10^{-5}) and η (typically 10^{-3}) are small.

The *gain* of the i-photon state Q_i is given by

$$Q_i = Y_i\frac{\mu^i}{i!}e^{-\mu}. \tag{8}$$

The gain Q_i is the product of the probability of Alice sending out an i-photon state (follows a Poisson distribution) and the conditional probability that Alice's i-photon state (and background) will lead to a detection event for Bob.

E. Quantum bit error rate

The error rate of the i-photon state, e_i, is given by

$$e_i = \frac{e_0Y_0 + e_{detector}\eta_i}{Y_i} \tag{9}$$

where $e_{detector}$ is the probability that a photon hit the erroneous detector. $e_{detector}$ characterizes the alignment and stability of the optical system. Experimentally, even at distances as long as 122 km, $e_{detector}$ is more or less independent of the distance. In what follows, we will assume that $e_{detector}$ is a constant. We will assume that the background is random. Thus the error rate of the background is $e_0=\frac{1}{2}$. Note that Eqs. (6)–(9) are satisfied for all $i=0,1,2,\ldots$. The overall gain is given by

$$Q_\mu = \sum_{i=0}^{\infty} Y_i\frac{\mu^i}{i!}e^{-\mu} = Y_0 + 1 - e^{-\eta\mu}. \tag{10}$$

The overall QBER is given by

$$E_\mu Q_\mu = \sum_{i=0}^{\infty} e_iY_i\frac{\mu^i}{i!}e^{-\mu} = e_0Y_0 + e_{detector}(1 - e^{-\eta\mu}). \tag{11}$$

III. PRACTICAL DECOY METHOD

In this section, we will first discuss the choice of μ for the signal state to maximize the key generation rate as given by Eq. (1). Then, we will consider a specific protocol of two weak decoy states and show how they can be used to estimate Y_1 and e_1 rather accurately. After that, we will show how to choose two decoy states to optimize the key generation rate in Eq. (1). As a whole, we have a practical decoy state protocol with two weak decoy states.

A. Choose optimal μ

Here we will discuss how to choose the expected photon number of signal states μ to maximize the key generation rate in Eq. (1).

Let us begin with a general discussion. On one hand, we need to maximize the gain of the single-photon state Q_1, which is the only source for the final secure key. To achieve this, heuristically, we should maximize the probability of Alice sending out single-photon signals. With a Poisson distribution of the photon number, the single-photon fraction in the signal source reaches its maximum when $\mu=1$. On the other hand, we have to control the gain of multiphoton states to ensure the security of the system. Thus, we should keep the fraction Q_1/Q_μ high, which requires μ not to be too large. Therefore, intuitively we have

$$\mu \in (0,1].$$

As will be noted in Sec. III B, Alice and Bob can estimate e_1 and Y_1 rather accurately in a simple decoy state protocol (e.g., one involving only two decoy states). Therefore, for ease of discussion, we will discuss the case where Alice and Bob can estimate e_1 and Y_1 perfectly. Minor errors in Alice's and Bob's estimations of e_1 and Y_1 will generally lead to rather modest change in the final key generation rate R. According to Eqs. (8) and (9), Q_1 will be maximized when $\mu=1$ and e_1 is independent of μ, so we can expect that the optimal expected photon number of the signal state is $\mu=O(1)$.

We consider the case where the background rate is low ($Y_0 \ll \eta$) and the transmittance is small $\eta \ll 1$ (typical values

TABLE I. Key parameters for QKD experiments.

Experiment	λ (nm)	α (dB/km)	$e_{detector}$ (%)	Y_0	η_{Bob}	f (MHz)
GYS [5]	1550	0.21	3.3	1.7×10^{-6}	0.045	2
Bourennane *et al.* [18]	1550	0.2	1	4×10^{-4}	0.143	0.1

are $Y_0=10^{-5}$ and $\eta=10^{-3}$). By substituting Eqs. (8), (9), (10), and (11) into Eq. (1), the key generation rate is given by

$$R \approx -\eta\mu f(e_{detector})H_2(e_{detector}) + \eta\mu e^{-\mu}[1-H_2(e_{detector})].$$

This rate is optimized if we choose $\mu=\mu_{optimal}$ which satisfies

$$(1-\mu)\exp(-\mu) = \frac{f(e_{detector})H_2(e_{detector})}{1-H_2(e_{detector})}, \tag{12}$$

where $e_{detector}$ is the probability that a photon hits the erroneous detector. Then, using the data shown in Table I extracted from a recent experiment [5], we can solve this equation and obtain that $\mu_{optimal}^{GYS}\approx 0.54$ for $f(e)=1$ and $\mu_{optimal}^{GYS}\approx 0.48$ for $f(e)=1.22$. As noted in [9], the key generation rate and distance are pretty stable against even a 20% change of μ.

B. General decoy method

Here we will give out the most general decoy state method with m decoy states. This extends our earlier work in [9].

Suppose Alice and Bob choose the signal and decoy states with expected photon numbers $\mu, \nu_1, \nu_2, \ldots, \nu_m$; they will get the following gains and QBER's for the signal state and decoy states:

$$Q_\mu e^\mu = \sum_{i=0}^{\infty} Y_i \frac{\mu^i}{i!},$$

$$E_\mu Q_\mu e^\mu = \sum_{i=0}^{\infty} e_i Y_i \frac{\mu^i}{i!},$$

$$Q_{\nu_1} e^{\nu_1} = \sum_{i=0}^{\infty} Y_i \frac{\nu_1^i}{i!},$$

$$E_{\nu_1} Q_{\nu_1} e^{\nu_1} = \sum_{i=0}^{\infty} e_i Y_i \frac{\nu_1^i}{i!},$$

$$Q_{\nu_2} e^{\nu_2} = \sum_{i=0}^{\infty} Y_i \frac{\nu_2^i}{i!},$$

$$E_{\nu_2} Q_{\nu_2} e^{\nu_2} = \sum_{i=0}^{\infty} e_i Y_i \frac{\nu_2^i}{i!},$$

$$\ldots$$

$$Q_{\nu_m} e^{\nu_m} = \sum_{i=0}^{\infty} Y_i \frac{\nu_m^i}{i!},$$

$$E_{\nu_m} Q_{\nu_m} e^{\nu_m} = \sum_{i=0}^{\infty} e_i Y_i \frac{\nu_m^i}{i!}. \tag{13}$$

Question. Given Eqs. (13), how can one find a tight lower bound on R, which is given by Eq. (1)? This is a main optimization problem for the design of decoy state protocols.

Note that in Eq. (1), the first term and q are independent of $\{Y_i\}$ and $\{e_i\}$. Combining with Eq. (8), we can simplify the problem: How can we lower-bound

$$P = Y_1[1-H_2(e_1)] \tag{14}$$

with the constraints given by Eqs. (13)?

When $m\to\infty$, Alice and Bob can solve all $\{Y_i\}$ and $\{e_i\}$ accurately in principle. This is the asymptotic case given in [9].

C. Two decoy states

As emphasized in [9], only a few decoy states are needed for practical implementations. A simple way to lower-bound Eq. (14) is to lower-bound Y_1 and upper-bound e_1. Intuitively, only two decoy states are needed for the estimation of Y_1 and e_1 and, therefore, for practical decoy state implementation. Here, we present a rigorous analysis to show more precisely how to use two weak decoy states to estimate the lower bound Y_1 and upper bound e_1.

Suppose Alice and Bob choose two decoy states with expected photon numbers ν_1 and ν_2 which satisfy

$$0 \leqslant \nu_2 \leqslant \nu_1,$$

$$\nu_1 + \nu_2 < \mu, \tag{15}$$

where μ is the expected photon number of the signal state.

1. Lower bound of Y_1

Similar to Eq. (10), the gains of these two decoy states are given by

$$Q_{\nu_1} = \sum_{i=0}^{\infty} Y_i \frac{\nu_1^i}{i!} e^{-\nu_1}, \tag{16}$$

$$Q_{\nu_2} = \sum_{i=0}^{\infty} Y_i \frac{\nu_2^i}{i!} e^{-\nu_2}. \tag{17}$$

First Alice and Bob can estimate the lower bound of background rate Y_0 by $e_{\nu 2}\times$ [Eq. (17)] $-e_{\nu 1}\times$ [Eq. (16)],

$$\begin{aligned}\nu_1 Q_{\nu_2} e^{\nu_2} - \nu_2 Q_{\nu_1} e^{\nu_1} &= (\nu_1 - \nu_2) Y_0 - \nu_1 \nu_2 \left(Y_2 \frac{\nu_1 - \nu_2}{2!} + Y_3 \frac{\nu_1^2 - \nu_2^2}{3!} + \cdots \right) \\ &\leqslant (\nu_1 - \nu_2) Y_0 .\end{aligned}$$

Thus, a crude lower bound of Y_0 is given by

$$Y_0 \geqslant Y_0^L = \max\left\{ \frac{\nu_1 Q_{\nu_2} e^{\nu_2} - \nu_2 Q_{\nu_1} e^{\nu_1}}{\nu_1 - \nu_2}, 0 \right\}, \tag{18}$$

where the equality sign will hold when $\nu_2 = 0$; that is to say, when a vacuum decoy ($\nu_2 = 0$) is used, Eq. (18) is tight.

Now, from Eq. (10), the contribution from multiphoton states (with photon number $\geqslant 2$) in the signal state can be expressed by

$$\sum_{i=2}^{\infty} Y_i \frac{\mu^i}{i!} = Q_\mu e^\mu - Y_0 - Y_1 \mu . \tag{19}$$

Combining Eqs. (16) and (17), under condition Eq. (15), we have

$$\begin{aligned} Q_{\nu_1} e^{\nu_1} - Q_{\nu_2} e^{\nu_2} &= Y_1(\nu_1 - \nu_2) + \sum_{i=2}^{\infty} \frac{Y_i}{i!} (\nu_1^i - \nu_2^i) \\ &\leqslant Y_1(\nu_1 - \nu_2) + \frac{\nu_1^2 - \nu_2^2}{\mu^2} \sum_{i=2}^{\infty} Y_i \frac{\mu^i}{i!} \\ &= Y_1(\nu_1 - \nu_2) + \frac{\nu_1^2 - \nu_2^2}{\mu^2} (Q_\mu e^\mu - Y_0 - Y_1 \mu) \\ &\leqslant Y_1(\nu_1 - \nu_2) + \frac{\nu_1^2 - \nu_2^2}{\mu^2} (Q_\mu e^\mu - Y_0^L - Y_1 \mu), \end{aligned} \tag{20}$$

where Y_0^L was defined in Eq. (18). Here, to prove the first inequality in Eq. (20), we have made use of the inequality that $a^i - b^i \leqslant a^2 - b^2$ whenever $0 < a + b < 1$ and $i \geqslant 2$. The equality sign holds for the first inequality in Eq. (20) if and only if Eve raises the yield of two-photon states and blocks all the states with photon number greater than 2 (this was also mentioned in [8]). The second equality in Eq. (20) is due to Eq. (18).

By solving inequality (20), the lower bound of Y_1 is given by

$$Y_1 \geqslant Y_1^{L,\nu_1,\nu_2} = \frac{\mu}{\mu\nu_1 - \mu\nu_2 - \nu_1^2 + \nu_2^2} \left(Q_{\nu_1} e^{\nu_1} - Q_{\nu_2} e^{\nu_2} - \frac{\nu_1^2 - \nu_2^2}{\mu^2} (Q_\mu e^\mu - Y_0^L) \right). \tag{21}$$

Then the gain of single-photon states is given by, according to Eq. (8),

$$Q_1 \geqslant Q_1^{L,\nu_1,\nu_2} = \frac{\mu^2 e^{-\mu}}{\mu\nu_1 - \mu\nu_2 - \nu_1^2 + \nu_2^2} \left(Q_{\nu_1} e^{\nu_1} - Q_{\nu_2} e^{\nu_2} - \frac{\nu_1^2 - \nu_2^2}{\mu^2} (Q_\mu e^\mu - Y_0^L) \right), \tag{22}$$

where Y_0^L is given by Eq. (18).

2. Upper bound of e_1

According to Eq. (11), the QBER of the weak decoy state is given by

$$E_{\nu_1} Q_{\nu_1} e^{\nu_1} = e_0 Y_0 + e_1 \nu_1 Y_1 + \sum_{i=2}^{\infty} e_i Y_i \frac{\nu_1^i}{i!}, \tag{23}$$

$$E_{\nu_2} Q_{\nu_2} e^{\nu_2} = e_0 Y_0 + e_1 \nu_2 Y_1 + \sum_{i=2}^{\infty} e_i Y_i \frac{\nu_2^i}{i!}. \tag{24}$$

An upper bound of e_1 can be obtained directly from Eqs. (23) and (24),

$$e_1 \leqslant e_1^{U,\nu_1,\nu_2} = \frac{E_{\nu_1} Q_{\nu_1} e^{\nu_1} - E_{\nu_2} Q_{\nu_2} e^{\nu_2}}{(\nu_1 - \nu_2) Y_1^{L,\nu_1,\nu_2}}. \tag{25}$$

Note that Alice and Bob should substitute the lower bound of Y_1, Eq. (21), into Eq. (25) to get an upper bound of e_1.

In summary, by using two weak decoy states that satisfy Eq. (15), Alice and Bob can obtain a lower bound for the yield Y_1 with Eq. (21) [and then the gain Q_1 with Eq. (22)] and an upper bound for the QBER e_1 with Eq. (25) for the single-photon signals. Subsequently, they can use Eq. (1) to work out the key generation rate as

$$R \geqslant q\{ - Q_\mu f(E_\mu) H_2(E_\mu) + Q_1^{L,\nu_1,\nu_2} [1 - H_2(e_1^{U,\nu_1,\nu_2})] \}. \tag{26}$$

This is the main procedure of our two-decoy-state protocol.

Now, the next question is, how good are our bounds for Y_1 and e_1 for our two-decoy-state protocol? In what follows, we will examine the performance of our two-weak-decoy-state protocol by considering first the asymptotic case where both ν_1 and ν_2 tend to 0. We will show that our bounds for Y_1 and e_1 are tight in this asymptotic limit.

3. Asymptotic case

We will now take the limit $\nu_1 \to 0$ and $\nu_2 \to 0$. When $\nu_2 < \nu_1 \ll \mu = O(1)$, substituting Eqs. (10), (16), and (17) into Eq. (21), the lower bound of Y_1 becomes

$$\begin{aligned} Y_1^{L,0} &= Y_1^{L,\nu_1,\nu_2}\big|_{\nu_1 \to 0, \nu_2 \to 0} \\ &= \frac{\mu}{\mu\nu_1 - \mu\nu_2 - \nu_1^2 + \nu_2^2} (Q_{\nu_1} e^{\nu_1} - Q_{\nu_2} e^{\nu_2})\big|_{\nu_1 \to 0, \nu_2 \to 0} \\ &= \frac{\mu}{\mu - \nu_1 - \nu_2} \frac{1}{\nu_1 - \nu_2} \\ &\quad \times [(Y_0 + \eta\nu_1) e^{\nu_1} - (Y_0 + \eta\nu_2) e^{\nu_2}]\big|_{\nu_1 \to 0, \nu_2 \to 0} \end{aligned}$$

$$= Y_0 + \eta, \tag{27}$$

which matches the theoretical value $Y_1 \cong Y_0 + \eta$ from Eq. (7). Substituting Eqs. (11), (23), and (24) into Eq. (25), the upper bound of e_1 becomes

$$e_1^{U,0} = e_1^{U,\nu_1,\nu_2}\big|_{\nu_1\to 0,\nu_2\to 0} = \frac{e_0 Y_0 + e_{detector}\eta}{Y_1}, \tag{28}$$

which matches the theoretical value from Eq. (9).

The above calculation seems to suggest that our two-decoy-state protocol is as good as the most general protocol in the limit $\nu_1, \nu_2 \to 0$. However, in real life, at least one of the two quantities ν_1 and ν_2 must take on a nonzero value. Therefore, we need to study the effects of finite ν_1 and ν_2. This will be our next subject.

4. Deviation from theoretical values

Here, we consider how finite values of ν_1 and perhaps ν_2 will change our bounds for Y_1 and e_1.

The relative deviation of Y_1 is given by

$$\beta_{Y1} = \frac{Y_1^{L,0} - Y_1^{L,\nu_1,\nu_2}}{Y_1^{L,0}}, \tag{29}$$

where $Y_1^{L,0}$ is the theoretical value of Y_1 given in Eqs. (7) and (27), and Y_1^{L,ν_1,ν_2} is an estimation value of Y_1 by our two-decoy-state method as given in Eq. (21).

The relative deviation of e_1 is given by

$$\beta_{e1} = \frac{e_1^{U,\nu_1,\nu_2} - e_1^{U,0}}{e_1^{U,0}}, \tag{30}$$

where $e_1^{U,0}$ is the theoretical value of e_1 given in Eqs. (9) and (28), and e_1^{U,ν_1,ν_2} is the estimation value of e_1 by our two-decoy-state method as given in Eq. (25).

Under the approximation $\eta \ll 1$ and taking the first order in ν_1 and ν_2, and substituting Eqs. (7), (10), (16)–(18), and (21) into Eq. (29), the deviation of the lower bound of Y_1 is given by

$$\begin{aligned} Y_1\beta_{Y1} &= Y_1^{L,0} - Y_1^{L,\nu_1,\nu_2} \\ &= Y_0 + \eta - \frac{\mu}{\mu\nu_1 - \mu\nu_2 - \nu_1^2 + \nu_2^2} \\ &\quad \times \left(Q_{\nu_1} e^{\nu_1} - Q_{\nu_2} e^{\nu_2} - \frac{\nu_1^2 - \nu_2^2}{\mu^2}(Q_\mu e^\mu - Y_0^L) \right) \\ &\approx \left(e^\mu - 1 - \mu - \frac{\mu^2}{2} \right)\left(\frac{1}{\mu - \nu_1 - \nu_2} - \frac{1}{\mu} \right) Y_0 \\ &\quad + (e^\mu - 1 - \mu)\frac{\nu_1 + \nu_2}{\mu - \nu_1 - \nu_2}\eta. \end{aligned} \tag{31}$$

Substituting Eqs. (9), (11), and (23)–(25), and (31) into Eq. (30), the deviation of the upper bound of e_1 is given by

$$e_1\beta_{e1} = e_1^{U,\nu,0} - e_1^{U,0} = e_1\beta_{Y1} + (\nu_1 + \nu_2)\left(e_1 - \frac{e_0 Y_0}{2Y_1} \right). \tag{32}$$

Now, from Eqs. (31) and (32), we can see that decreasing $\nu_1 + \nu_2$ will improve the estimation of Y_1 and e_1. So, the smaller $\nu_1 + \nu_2$ is, the higher the key generation rate R is. In the Appendix, we will prove that decreasing $\nu_1 + \nu_2$ will improve the estimation of Y_1 and e_1 in a general sense (i.e., without the limit $\eta \ll 1$ and taking the first order in ν_1 and ν_2). Therefore, we have reached the following important conclusion: for any fixed value of ν_1, the choice $\nu_2 = 0$ will optimize the key generation rate. In this sense, the vacuum + weak decoy state protocol, as first proposed in an intuitive manner in [9], is, in fact, optimal.

The above conclusion highlights the importance of the vacuum + weak decoy state protocol. We will discuss it in Sec. III D. Nonetheless, as remarked earlier, in practice, it might not be easy to prepare a true vacuum state (with, say, VOAs). Therefore, our general theory on nonzero decoy states, presented here, is important.

D. Vacuum + weak decoy state

Here we will introduce a special case of Sec. III C with two decoy states: vacuum and a weak decoy state. This special case was first proposed in [9] and analyzed in [13]. In the end of Sec. III C, we pointed out that this case is optimal for the two-decoy-state method.

1. Vacuum decoy state

Alice shuts off her photon source to perform the vacuum decoy state. Through this decoy state, Alice and Bob can estimate the background rate,

$$Q_{vacuum} = Y_0,$$

$$E_{vacuum} = e_0 = \frac{1}{2}. \tag{33}$$

The dark counts occur randomly; thus the error rate of the dark count is $e_0 = \frac{1}{2}$.

2. Weak decoy state

Alice and Bob choose a relatively weak decoy state with expected photon number $\nu < \mu$.

Here is the key difference between this special case and our general case of the two-decoy-state protocol. Now, from the vacuum decoy state Eq. (33), Alice and Bob can estimate Y_0 accurately. So, the second inequality of Eq. (20) will be tight. As in Eq. (21), the lower bound of Y_1 is given by

$$\begin{aligned} Y_1 &\geqslant Y_1^{L,\nu,0} = Y_1^{L,\nu,\nu_2}\big|_{\nu_2\to 0} \\ &= \frac{\mu}{\mu\nu - \nu^2}\left(Q_\nu e^\nu - Q_\mu e^\mu \frac{\nu^2}{\mu^2} - \frac{\mu^2 - \nu^2}{\mu^2} Y_0 \right). \end{aligned} \tag{34}$$

So the gain of a single-photon state is given by [Eq. (8)]

$$Q_1 \geqslant Q_1^{L,\nu,0} = \frac{\mu^2 e^{-\mu}}{\mu\nu - \nu^2}\left(Q_\nu e^\nu - Q_\mu e^\mu \frac{\nu^2}{\mu^2} - \frac{\mu^2 - \nu^2}{\mu^2} Y_0 \right). \tag{35}$$

We remark that Eq. (34) can be used to provide a simple derivation of the fraction of "tagged photons" Δ found in Wang's paper [13],

$$\Delta = \frac{Q_\nu - Y_0 e^{-\nu} - Y_1 \nu e^{-\nu}}{Q_\nu} \leqslant \frac{Q_\nu - Y_0 e^{-\nu} - Y_1^{L,\nu,0} \nu e^{-\nu}}{Q_\nu}$$
$$= \frac{Q_\nu - Y_0 e^{-\nu} - [\mu e^{-\nu}/(\mu-\nu)]\{Q_\nu e^{\nu} - Q_\mu e^{\mu(\nu 2/\mu^2)} - [(\mu^2-\nu^2)/\mu^2]Y_0\}}{Q_\nu} = \frac{\nu}{\mu-\nu}\left(\frac{\nu e^{-\nu} Q_\mu}{\mu e^{-\mu} Q_\nu} - 1\right) + \frac{\nu e^{-\nu} Y_0}{\mu Q_\nu}. \quad (36)$$

Indeed, if we replace ν by μ and μ by μ', Eq. (36) will be exactly the same as Eq. (2).

According to Eq. (25), the upper bound of e_1 is given by

$$e_1 \leqslant e_1^{U,\nu,0} = \frac{E_\nu Q_\nu e^{\nu} - e_0 Y_0}{Y_1^{L,\nu,0}\nu}. \quad (37)$$

3. Deviation from theoretical values

Considering the approximation $\eta \ll 1$ and taking the first order in ν, as in Eqs. (31) and (32), the theoretical deviations of the vacuum+weak decoy method are given by

$$Y_1\beta_{Y1} = Y_1^{L,0} - Y_1^{L,\nu,0}$$
$$= Y_0 + \eta - \frac{\mu}{\mu\nu-\nu^2}\left(Q_\nu e^{\nu} - Q_\mu e^{\mu}\frac{\nu^2}{\mu^2} - \frac{\mu^2-\nu^2}{\mu^2}Y_0\right)$$
$$\approx \frac{\nu}{\mu}(e^{\mu} - 1 - \mu)\eta + \frac{\nu}{\mu^2}\left(e^{\mu} - 1 - \mu - \frac{\mu^2}{2}\right)Y_0,$$

$$e_1\beta_{e1} = e_1^{U,\nu,0} - e_1^{U,0} \approx e_1\beta_{Y1} + \nu\left(e_1 - \frac{e_0 Y_0}{2Y_1}\right),$$

from which we can see that decreasing ν will improve the estimation of Y_1 and e_1. So, the smaller ν is, the higher the key generation rate R is. Later, in Sec. IV, we will take into account statistical fluctuations and give an estimation of the optimal value of ν which maximizes the key generation rate.

E. One decoy state

Here we will discuss a decoy state protocol with only one decoy state. Such a protocol is easy to implement in experiments, but may generally not be optimal. As noted earlier, we have successfully performed an experimental implementation of one-decoy-state QKD in [15].

1. A simple proposal

A simple method to analyze one-decoy-state QKD is by substituting an upper bound of Y_0 into Eq. (34) and a lower bound of Y_0 into Eq. (37) to lower-bound Y_1 and upper-bound e_1.

An upper bound of Y_0 can be derived from Eq. (11),

$$Y_0 \leqslant \frac{E_\mu Q_\mu e^{\mu}}{e_0}. \quad (38)$$

Substituting the above upper bound into Eq. (34), we get a lower bound on Y_1,

$$Y_1 \geqslant \bar{Y}_1^{L,\nu} = \frac{\mu}{\mu\nu-\nu^2}\left(Q_\nu e^{\nu} - Q_\mu e^{\mu}\frac{\nu^2}{\mu^2} - E_\mu Q_\mu e^{\mu}\frac{\mu^2-\nu^2}{e_0\mu^2}\right). \quad (39)$$

A simple lower bound on e_1 can be derived as follows:

$$e_1 \leqslant \bar{e}_1^{U,\nu} = \frac{E_\mu Q_\mu e^{\mu}}{Y_1^{L,\mu,0}\mu}. \quad (40)$$

Now, by substituting Eqs. (39) and (40) into Eq. (1), one obtains a simple lower bound of the key generation rate. The above lower bound has recently been used in our experimental decoy state QKD paper [15]. (In our experimental decoy QKD paper [15], we simplify our notation by denoting $\bar{Y}_1^{L,\nu}$ by simply Y_1^L and $\bar{e}_1^{U,\nu}$ by e_1^U.)

2. Tighter bound

Another method is to apply the results of the vacuum+weak decoy described in Sec. III D.

Let us assume that Alice and Bob perform the vacuum+weak decoy method, but they prepare very few states as the vacuum state. So they cannot estimate Y_0 very well. We claim that a single decoy protocol is the same as a vacuum+weak decoy protocol, except that we do not know the value of Y_0. Since Alice and Bob do not know Y_0, Eve can pick Y_0 as she wishes. We argue that, on physical ground, it is advantageous for Eve to pick Y_0 to be zero. This is because Eve may gather more information on the single-photon signal than the vacuum. Therefore, the bound for the case $Y_0=0$ should still apply to our one-decoy protocol. [We have explicitly checked mathematically that our following conclusion is correct, after lower-bounding Eq. (14) directly.] For this reason, Alice and Bob can derive a bound on the key generation rate R by substituting the following values of Y_1^{trial} and e_1^{trial} into Eq. (1):

$$Y_1^{trial} = \frac{\mu}{\mu\nu-\nu^2}\left(Q_\nu e^{\nu} - Q_\mu e^{\mu}\frac{\nu^2}{\mu^2}\right),$$

$$e_1^{trial} = \frac{E_\nu Q_\nu e^{\nu}}{Y_1^{trial}\nu}. \quad (41)$$

F. Example

Let us return to the two-decoy-state protocol. In Eqs. (27) and (28), we have shown that the two-decoy-state method is optimal in the asymptotic case where $\nu_1, \nu_2 \to 0$, in the sense that its key generation rate approaches the most general de-

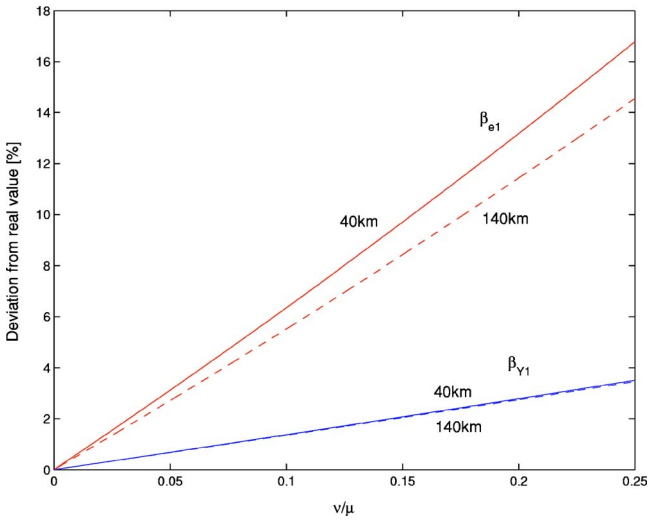

FIG. 1. (Color online) The solid lines show the relative deviations of Y_1^{L,ν_1,ν_2} and e_1^{U,ν_1,ν_2} from the asymptotic values (i.e., the case $\nu_1,\nu_2\to 0$) as functions of ν/μ (where $\nu=\nu_1$) with the fiber length 40 km, and the dashed lines show the case of 140 km. The bounds Y_1^{L,ν_1,ν_2} and e_1^{U,ν_1,ν_2} are given by Eqs. (34) and (37), and the true values are given by Eqs. (7) and (9). We consider the vacuum+weak protocol here ($\nu_1=\nu$ and $\nu_2=0$). The expected photon number is $\mu=0.48$ as calculated from Eq. (12). The parameters used are from GYS [5] as listed in Table I.

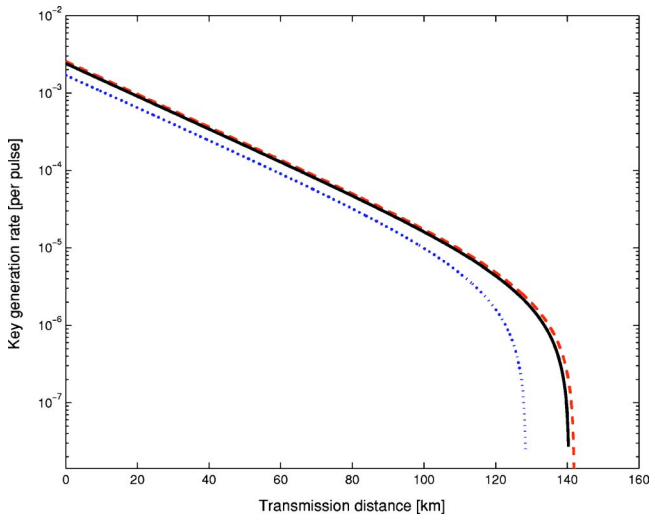

FIG. 2. (Color online) The dashed line shows the asymptotic decoy state method (with infinite number of decoy states) with a maximal secure distance of 142.05 km, using Eq. (1). The solid line shows our vacuum+weak decoy method, Eq. (42), with $\mu=0.48$, $\nu_1=0.05$, and $\nu_2=0$. It uses a strong version of the GLLP protocol and its maximal distance is 140.55 km. The dotted line shows the asymptotic case of Wang's decoy method, Eq. (43) with $\mu=0.30$. It uses a weak version of the GLLP protocol and its maximal distance is about 128.55 km. This shows that our vacuum+weak decoy protocol comes very close to the asymptotic limit and performs better than even the asymptotic case of Wang's decoy method. The data are from GYS [5] as listed in Table I.

coy state method of having an infinite number of decoy states. Here, we will give an example to show that, even in the case of finite ν_1 and ν_2, the performance of our two-decoy-state method is only slightly worse than that of the perfect decoy method. We will use the model in Sec. II to calculate the deviations of the estimated values of Y_1 and e_1 from our two-decoy-state method from the correct values. We use the data of GYS [5] with key parameters listed in Table I.

For simplicity, we will use a special two-decoy-state method: vacuum+weak. According to Eq. (12), the optimal expected photon number is $\mu=0.48$. We change the expected photon number of weak decoy ν to see how the estimates, described by Eqs. (34) and (37), deviate from the asymptotic values, Eqs. (7) and (9). The deviations are calculated by Eqs. (29) and (30). The results are shown in Fig. 1. From Fig. 1, we can see that the estimate for Y_1 is very good. Even at $\nu/\mu=25\%$, the deviation is only 3.5%. The estimate for e_1 is slightly worse. The deviation will go to 16.8% when $\nu/\mu=25\%$. The deviations do not change much with fiber length. Later, in Sec. IV, we will discuss how to choose optimal ν when statistical fluctuations due to a finite experimental time are taken into account.

Let R^L denote for the lower bound of key generation rate, according to Eq. (1),

$$R^L = q\{-Q_\mu f(E_\mu)H_2(E_\mu) + Q_1^{L,\nu,0}[1-H_2(e_1^{U,\nu,0})]\}, \tag{42}$$

where $q=\frac{1}{2}$ with the standard BB84 protocol. The parameters can be calculated from Eqs. (10), (11), (35), and (37) and use $f(e)=1.22$, which is the upper bound of $f(e)$ in secure distance for this experiment [12]. Equation (5) shows the relationship between η and distance. The results are shown in Fig. 2.

Now, from Fig. 2, we can see that even with finite ν (say, 0.05), the vacuum+weak protocol performs very like the asymptotic one.

We note that Wang [13] has also studied a decoy state protocol, first proposed by us [9], with only two decoy states for the special case where one of them is a vacuum. In [13] the second decoy state is used to estimate the multiphoton fraction Δ and use the formula directly from GLLP [7] to calculate the key generation rate by Eq. (3).

In Fig. 2, we compare the key generation rates of our two-decoy-state method and Wang's method [13] and find that our method performs better. In what follows, we compare the differences between our method and that of Wang.

(a) We consider error correction inefficiency $f(e)$ for practical protocols. Wang did not consider this real-life issue. For a fair comparison, we add this factor to Eq. (3):

$$R \geqslant qQ_\mu\left\{-f(E_\mu)H_2(E_\mu) + (1-\Delta)\left[1-H_2\left(\frac{E_\mu}{1-\Delta}\right)\right]\right\}. \tag{43}$$

(b) Apparently, the value of μ was chosen in [13] in an *ad hoc* manner, whereas we performed optimization in Sec. III A and found that for GYS, the optimal value of $\mu=0.48$ for our two-decoy-state method. Now, the best (asymptotic) estimate Wang's method can make is that Δ

$=\mu$ when $\mu' \to \mu$. For a fair comparison, we have performed an optimization of Wang's asymptotic result Eq. (43) as well (as in Sec. III A) and found that the value $\mu \approx 0.30$ optimizes the key generation rate in Wang's method.

(c) In Eqs. (27) and (28), we show that our two-decoy-state method approaches a fundamental limit of the decoy state (the infinite decoy state protocol) while the asymptotic result in Wang [13] is strictly bounded away from the fundamental limit. Even with a finite v_1, our vacuum+weak protocol is better than Wang's asymptotic case.

(d) Why do we get a stronger result than Wang [13]? Wang did not estimate e_1 and used $E_\mu/(1-\Delta)$ as the upper bound of e_1 (this corresponds to a weak version of the GLLP result [7]). We estimate e_1 more accurately following GLLP (a strong version of the GLLP result).

IV. STATISTICAL FLUCTUATIONS

In this section, we would like to discuss the effect of finite data-set size in real-life experiments on our estimation process for Y_1 and e_1. We will also discuss how statistical fluctuations might affect our choice of ν_1 and ν_2. We will provide a list of those fluctuations and discuss how we will deal with them. We remark that Wang [13] has previously considered the issue of fluctuations of Y_1.

All real-life experiments are done in a finite time. Ideally, we would like to consider a QKD experiment that can be performed within, say, a few hours or so. This means that our data-set size is finite. Here, we will see that this type of statistical fluctuation is a rather complex problem. We do not have a full solution to the problem. Nonetheless, we will provide some rough estimation based on standard error analysis which suggests that the statistical fluctuation problem of the two-decoy-state method for a QKD experiment appears to be under control, if we run an experiment over only a few hours.

A. What parameters are fluctuating?

Recall that from Eq. (1), there are four parameters that we need to take into account: the gain Q_μ and QBER E_μ of the signal state and the gain Q_1 and QBER e_1 of the single-photon state. The gain of the signal state Q_μ is measured directly from experiment. We note that the fluctuations of the the signal error rate E_μ are not important because E_μ is *not* used at all in the estimation of Y_1 and e_1. [See Eqs. (21) and (25) or Eqs. (35) and (37).] Therefore, the important issue is the statistical fluctuations of Q_1 and e_1 due to the finite data-set size of signal states and decoy states.

To show the complexity of the problem, we will now discuss the following five sources of fluctuations. The first thing to notice is that, in practice, the intensity of the lasers used by Alice will be fluctuating. In other words, even the parameters μ, ν_1, and ν_2 suffer from small statistical fluctuations. Without hard experimental data, it is difficult to pinpoint the extent of their fluctuations. To simplify our analysis, we will ignore their fluctuations in this paper.

The second thing to notice is that so far in our analysis we have assumed that the proportion of photon number eigenstates in each type of state is fixed. For instance, if N signal states of intensity μ are emitted, we assume that exactly $N\mu e^{-\mu}$ out of the N signal states are single photons. In real life, the number $\mu e^{-\mu}$ is only a probability; the actual number of single-photon signals will fluctuate statistically. The fluctuation here is dictated by the law of large number though. So this problem should be solvable. For simplicity, we will neglect this source of fluctuations in this paper. (It was subsequently pointed out to us by Gottesman and Preskill that the above two sources of fluctuations can be combined into the fluctuations in the photon number frequency distribution of the underlying signal and decoy states. These fluctuations will generally average out to zero in the limit of a large number of signals, provided that there is no systematic error in the experimental setup.)

The third thing to notice is that, as noted by Wang [13], the yield Y_i may fluctuate in the sense that Y_i for the signal state might be slightly different from Y_i' of the decoy state. We remark that if one uses the vacuum state as one of the decoy states, then by observing the yield of the vacuum decoy state, conceptually, one has a very good handle on the yield of the vacuum component of the signal state (in terms of hypergeometric functions). Note, however, that the background rate is generally rather low (typically 10^{-5}). So, to obtain a reasonable estimation on the background rate, a rather large number (say 10^7) of vacuum decoy states will be needed. (As noted in [9], even a 20% fluctuation in the background will have small effect on the key generation rates and distances.) Note that, with the exception of the case $n=0$ (the vacuum case), neither Y_i and Y_i' is directly observable in an experiment. In a real experiment, one can measure only some *averaged* properties. For instance, the yield Q_μ of the signal state, which can be experimentally measured, has its origin as the weighted averaged yields of the various photon number eigenstates Y_i whereas that for the decoy state is given by the weighted average of Y_i'. How to relate the observed averaged properties, e.g., Q_μ, to the underlying values of the Y_i's is a challenging question. In summary, owing to the fluctuations of Y_i for $n>0$, it is not clear to us how to derive a closed-form solution to the problem.

Fourth, we note that the error rates e_i for the signal can also be different from the error rates e_i for the decoy state, due to underlying statistical fluctuations. Actually, the fluctuation of e_1 appears to the dominant source of errors in the estimation process. (See, for example, Table II.) This is because the parameter e_1 is rather small (say a few percent) and it appears in combination with another small parameter Y_1 in Eq. (11) for the QBER.

Fifth, we noted that for security in the GLLP [7] formula [Eq. (1)], we need to correct phase errors, rather than bit-flip errors. From Shor and Preskill's proof [3], we know that the bit-flip error rate and the phase error rate are supposed to be the same only in the asymptotic limit. Therefore, for a finite data set, one has to consider statistical fluctuations. This problem is well studied [3]. Since the number of signal states is generally very big, we will ignore this fluctuation from now on.

Qualitatively, the yields of the signal and decoy states tend to decrease exponentially with distance. Therefore, statistical fluctuations tend to become more and more important

TABLE II. The pulse number distribution and ν are calculated from Eqs. (46) and (47). $\tilde{B}$ is the lower bound of final key bits. All results are obtained by numerical analysis using MATLAB. The variable $\tilde{\beta}_{Y1}$ denotes the relative error in our estimation process of Y_1 from its true value by using the data from a finite experiment. This relative error originates from statistical fluctuations. This definition contrasts with the definition of β_{Y1} in Eq. (29) which refers to the relative difference between the values of Y_1 for the case (i) where ν_1 and ν_2 are finite and the case (ii) where ν_1 and ν_2 approach zero. Similarly, other β's denote the relative errors in our estimates for the corresponding variables in the subscript of β. All the statistical fluctuation is of the confidence interval of ten standard deviations [i.e., $1-(1.5)\times10^{-23}$]. The data come from GYS [5], listed in Table I.

l	μ	u_α	N	N_S	N_1	N_2
103.62 km	0.479	10	6×10^9	3.98×10^9	1.76×10^9	2.52×10^8
η	ν	$\tilde{B}[bits]$	β_{Y0}	β_{Y1}	β_{e1}	β_R
3×10^{-4}	0.127	2.17×10^4	48.31%	7.09%	97.61%	74.11%

as the distance of QKD increases. In general, as the distance of QKD increases, larger and large data sets will be needed for the reliable estimation of Y_1 and e_1 (and hence R), thus requiring a longer QKD experiment.

In this paper, we will neglect the fluctuations due to the first two and the fifth sources listed above. Even though we cannot find any closed-form solution for the third and fourth sources of fluctuations, it should be possible to tackle the problem by simulations. Here, we are contented with a more elementary analysis. We will simply apply standard error analysis to perform a rough estimation on the effects of fluctuations due to the third and fourth sources. We remark that the origin of the problem is strictly classical statistical fluctuations. There is nothing quantum in this statistical analysis. While standard error analysis (using essentially normal distributions) may not give a completely correct answer, we expect that it is correct at least in the order of magnitude.

Our estimation, which will be presented below, shows that, for long-distance (>100 km) QKD with our two-decoy-state protocol, the statistical fluctuations effect (from the third and fourth sources only) appears to be manageable. This is so provided that a QKD experiment is run for a reasonable period of time of only a few hours. Our analysis supports the viewpoint that our two-decoy-state protocol is practical for real-life implementations.

We remark in passing that the actual classical memory space requirement for Alice and Bob is rather modest (<1 Gbyte) because at long distance, only a small fraction of the signals will give rise to detection events.

We emphasize that we have not fully solved the statistical fluctuation problem for decoy state QKD. This problem turns out to be quite complex. We remark that this statistical fluctuation problem will affect all earlier results including [8,9,13]. In future investigations, it will be interesting to study the issue of classical statistical fluctuations in more detail.

B. Standard error analysis

In what follows, we present a general procedure for studying the statistical fluctuations (due to the third and fourth sources noted above) by using standard error analysis.

Denote the number of pulses (sent by Alice) for the signal as N_S, and for two decoy states as N_1 and N_2. Then, the total number of pulses sent by Alice is given by

$$N = N_S + N_1 + N_2. \tag{44}$$

Then the parameter q in Eq. (1) is given by

$$q = \frac{N_S}{2N}. \tag{45}$$

Here we assume Alice and Bob perform the standard BB84 protocol. So there is a factor of $\frac{1}{2}$.

In practice, since N is finite, the statistical fluctuations of Q_1 and e_1 cannot be neglected. All these additional deviations will be related to data sizes N_S, N_1, and N_2 and can, in principle, be obtained from statistical analysis. A natural question to ask is the following. Given total data size N=const, how can we distribute it to N_S, N_1, and N_2 to maximize the key generation rate R ? This question also relates to another one: how can we choose optimal weak decoys ν_1 and ν_2 to minimize the effects of statistical fluctuations?

In principle, our optimization procedure should go as follows. First (this is the hard part), one needs to derive a lower bound of Q_1 and an upper bound of e_1 (as functions of data-set size N_S, N_1, N_2, ν_1, and ν_2), taking into full account the statistical fluctuations. Second, one substitutes those bounds in Eq. (1) to calculate the lower bound of the key generation rate, denoted by R^L. Thus, R^L is a function of N_S, N_1, N_2, ν_1, and ν_2, and will be maximized when the optimal distribution satisfies

$$\frac{\partial R^L}{\partial N_S} = \frac{\partial R^L}{\partial N_1} = \frac{\partial R^L}{\partial N_2} = 0, \tag{46}$$

given $N=N_S+N_1+N_2=$const.

C. Choice of ν_1 and ν_2

Now, from the theoretical deviations of Y_1 and e_1, Eqs. (29) and (30), reducing ν may decrease the theoretical deviations. We need to take statistical fluctuations into account. Given a fixed N_1+N_2, reducing ν_1 and ν_2 will decrease the

number of detection events of decoy states, which in turn causes a larger statistical fluctuation. Thus, there exists an optimal choice of ν_1 and ν_2 that maximizes the lower bound of the key generation rate R^L,

$$\frac{\partial R^L}{\partial \nu_1} = \frac{\partial R^L}{\partial \nu_2} = 0,$$

which can be simplified to

$$\frac{\partial}{\partial \nu_1}\{\hat{Y}_1^{L,\nu_1,\nu_2}[1 - H_2(\hat{e}_1^{U,\nu_1,\nu_2})]\} = 0,$$

$$\frac{\partial}{\partial \nu_2}\{\hat{Y}_1^{L,\nu_1,\nu_2}[1 - H_2(\hat{e}_1^{U,\nu_1,\nu_2})]\} = 0, \tag{47}$$

where $\hat{Y}_1^{L,\nu_1,\nu_2}$ and $\hat{e}_1^{U,\nu_1,\nu_2}$ are the lower bound to Y_1 and the upper bound to e_1 when statistical fluctuations are considered.

Given the total data-set size in Eq. (44), in principle, one can solve Eqs. (46) and (47) to get N_S, N_1, N_2, ν_1, and ν_2.

D. Simulation

In real life, solving Eqs. (46) and (47) is a complicated problem. In what follows, we will be contented with a rough estimation procedure using standard error analysis commonly used by experimentalists.

Some assumptions. In the following, we will discuss the vacuum+weak decoy method only.

(1) The signal state is used much more often than the two decoy states. Given the large number of signal states, it is reasonable to ignore the statistical fluctuations in signal states.

(2) We assume that the decoy state used in the actual experiment is conceptually only a part of an infinite population of decoy states. There are underlying values for Q_ν and E_ν as defined by the population of decoy states. In each realization, the decoy state allows us to obtain some estimates for these underlying Q_ν and E_ν. Alice and Bob can use the fluctuations of Q_ν, E_ν to calculate the fluctuation of the estimates of Y_1 and e_1.

(3) We neglect the change of $f(E_\mu)$ due to small changes in E_μ.

(4) When the number of events (e.g., the total detection event of the vacuum decoy state) is large (say >50), we assume that the statistical characteristic of a parameter can be described by a *normal* distribution.

We will use the experiment parameters in Table I and show numerical solutions of Eqs. (44), (46), and (47). We pick the total data size to be $N=6\times10^9$. Now, the GYS experiment [5] has a repetition rate of 2 MHz and an up time of less than 50% [20]. Therefore, it should take only a few hours to perform our proposed experiment. The optimal $\mu=0.48$ can be calculated by Eq. (12) and we use $f(e)=1.22$.

In the fiber length of 103.62 km ($\eta=3\times10^{-4}$), the optimal pulse distribution of data, ν, and the deviations from the perfect decoy method are listed in Table II.

For each fiber length we can solve Eqs. (46) and (47) to get N_S, N_E, N_1, N_2, and ν.

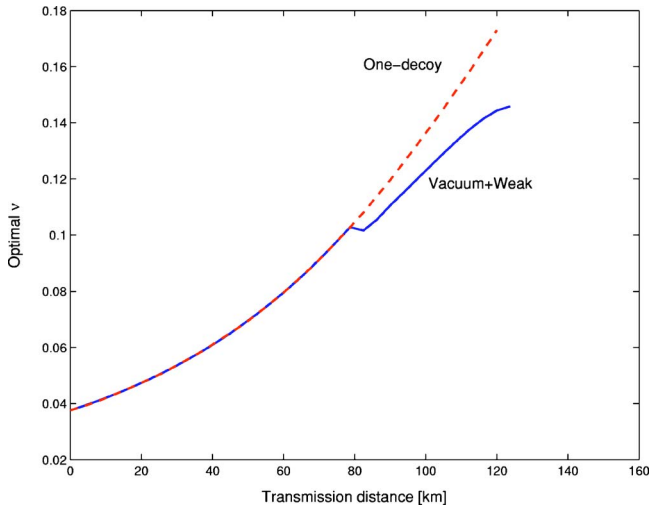

FIG. 3. (Color online) The solid line shows the simulation result of the vacuum+weak protocol [Eqs. (34) and (37)] with statistical fluctuations. The dashed line shows the result for the one-decoy-state method [Eq. (41)]. Here, we pick the data-set size (total number of pulses emitted by Alice) to be $N=6\times10^9$. We find the optimal ν for each fiber length by numerically solving Eqs. (44), (46), and (47). The confidence interval for statistical fluctuation is ten standard deviations [i.e., $1-(1.5)\times10^{-23}$]. The data are from GYS [5] as listed in Table I. The expected photon number of the signal state is calculated by Eq. (12), getting $\mu=0.48$. The second decoy state (vacuum decoy) becomes useful at 82 km.

Figure 3 shows how the optimal ν changes with fiber length. We can see that the optimal ν is small (~0.1) through the whole distance. In fact, it starts at a value $\nu\approx0.04$ at zero distance and increases almost linearly with the distance.

Figure 4 shows the vacuum+weak state with statistical fluctuations as compared to the asymptotic case of an infinite decoy state and without statistical fluctuations. We can see that even taking into account the statistical fluctuations, the vacuum+weak protocol is not far from the asymptotic result. In particular, in the short-distance region, our two-decoy-state method with statistical fluctuations approaches the performance of the asymptotic limit of infinite decoy states and no statistical fluctuations. This is so because the channel is not very lossy and statistical fluctuations are easily under control. This fact highlights the feasibility of our proposal.

Wang [13] picked the total data-set size $N=8.4\times10^{10}$. For long-distance QKD, this will take more than one day of experiment with the current GYS setup [5]. In order to perform a fair comparison with Wang result [13], we will now use the data-set size $N=8.4\times10^{10}$. Figure 5 shows R^L vs fiber length l with $N=8.4\times10^{10}$ fixed and compares our vacuum+weak protocol with Wang's result.

We have the following comments.

(a) Wang [13] chooses the value of μ in an *ad hoc* manner. Here we note that, for Wang's asymptotic case, the optimal choice of μ is $\mu\in[0.25,0.3]$.

(b) Even if we choose $\mu\in[0.25,0.3]$, the maximal secure distance of Wang's asymptotic case is still less than that of our two-decoy-state method with statistical fluctuations. In other words, the performance of our two-decoy-state method with statistical fluctuations is still better than the asymptotic

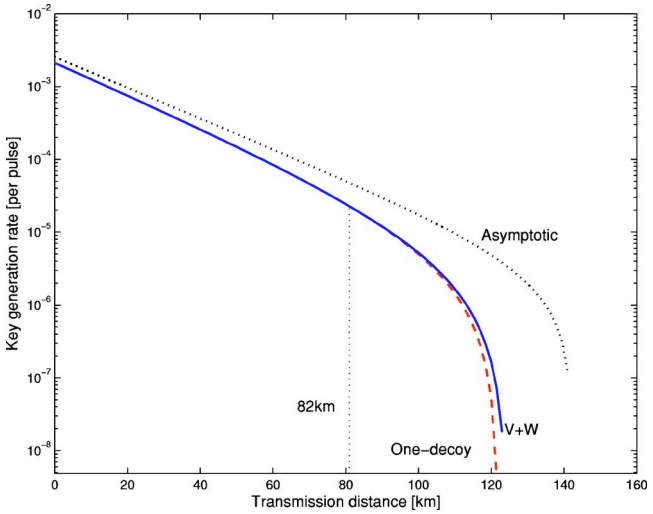

FIG. 4. (Color online) The dotted line shows the performance of the perfect decoy state method (with infinite number of decoy states and no statistical fluctuations). The maximal distance is about 142 km. The solid line shows the simulation result of the vacuum+weak protocol [Eqs. (34) and (37)] with statistical fluctuations. Its maximal distance is about 125 km. The dashed line shows the result for the one-decoy-state method [Eq. (41)] with maximal distance 122 km. We pick a data-set size (i.e., total number of pulses emitted by Alice) to be $N=6\times10^9$. Note that even with statistical fluctuations and a rather modest data-set size, our vacuum+weak decoy protocol comes rather close to asymptotic limit, particularly at short distances. The second decoy state (vacuum decoy) becomes useful at 82 km. The data are from GYS [5] as listed in Table I. The expected photon number of the signal state is calculated by Eq. (12), getting $\mu=0.48$.

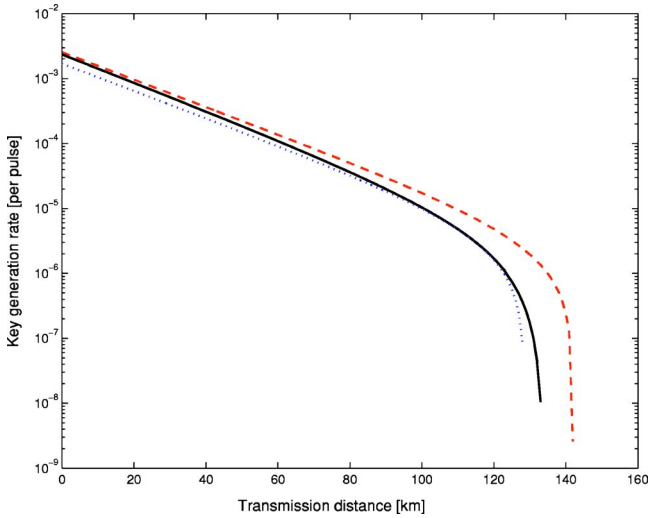

FIG. 5. (Color online) Here, we consider the data-set size (i.e., the number of pulses emitted by Alice) to be $N=8.4\times10^{10}$, following Wang [13]. The dashed line shows the performance of the perfect decoy state method. Its maximal distance is 142 km. The solid line shows the simulation result of the vacuum+weak decoy state method with statistical fluctuations. Its maximal distance is 132 km. The dotted line shows the asymptotic case (i.e., an idealized version) of Wang's method. Its maximal distance is 128.55 km. This figure shows clearly that with a data-set size $N=8.4\times10^{10}$, our protocol, which considers statistical fluctuations, performs better even than the idealized version of Wang's protocol, where statistical fluctuations are neglected. For our asymptotic case and two decoy state with statistical fluctuation $\mu=0.48$, and for Wang's asymptotic case $\mu=0.3$, which are optimized.

value (i.e., without considering statistical fluctuations) given by Wang's method.

(c) Note that the GYS experiment [5] has a very low background rate ($Y_0=1.7\times10^{-6}$) and high $e_{detector}$. The typical values of these two key parameters are $Y_0=10^{-5}$ and $e_{detector}=1\%$. If the background rate is higher and $e_{detector}$ is lower, then our results will have more advantage over Wang's. We illustrate this fact in Fig. 6 by using the data from the Bourennane *et al.* experiment [19].

V. CONCLUSION

We studied the two-decoy-state protocol where two weak decoy states of intensities ν_1 and ν_2 and a signal state with intensity μ are employed. We derived a general formula for the key generation rate R of the protocol and showed that the asymptotically limiting case where ν_1 and ν_2 tend to zero gives an optimal key generation rate which is the same as having infinite number of decoy states. This result seems to suggest that there is no fundamental conceptual advantage in using more than two decoy states. Using the data from the GYS experiment [5], we studied the effect of finite ν_1 and ν_2 on the value of the key generation rate R. In particular, we considered a vacuum+weak protocol, proposed in [9] and analyzed in [13], where $\nu_2=0$, and showed that R does not change much even when ν_1/μ is as high as 25%. We also derived the optimal choice of expected photon number μ of the signal state, following our earlier work [9]. Finally, we considered the issue of statistical fluctuations due to a finite data-set size. We remark that statistical fluctuations have also been considered in the recent work of Wang [13]. Here, we listed five different sources of fluctuations. While the problem is highly complex, we provided an estimation based on standard error analysis. We believe that such an analysis, while not rigorous, will give at least the correct order of magnitude estimation to the problem. This is so because this is a classical estimation problem. There is nothing quantum about it. That is to say, there are no subtle quantum attacks to consider. Our estimation showed that two-decoy-state QKD appears to be highly practical. Using data from a recent experiment [5], we showed that, even for long-distance (i..e, over 100 km) QKD, only a few hours of data are sufficient for its implementation. The memory size requirement is also rather modest (<1 GByte). A caveat is that we have not considered the fluctuations of the laser intensities of Alice, i.e., the values of μ, ν_1, and ν_2. This is because we do not have reliable experimental data to perform such an investigation. For short-distance QKD, the effects of statistical fluctuations are suppressed because the transmittance and useful data rate are much higher than for long-distance QKD. Finally, we noted that statistical fluctuations will affect our choice of decoy states ν_1 and ν_2 and performed an optimization for the special case where $\nu_2=0$.

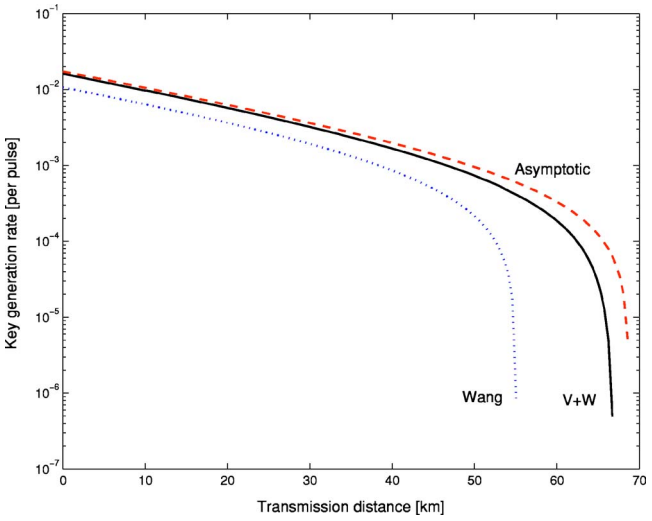

FIG. 6. (Color online) Here, we compare various protocols using the parameters of Bourenanne *et al.* [19], listed in Table I and [18]. The dashed line shows the performance of the perfect decoy state method. It has a maximal secure distance of about 68.6 km. The solid line shows the simulation result of the vacuum+weak decoy state method with statistical fluctuations. The maximal distance is about 67.2 km. The dotted line shows the asymptotic case (i.e., neglecting statistical fluctuations) of Wang's method, whose maximal distance is about 55.5 km. For our asymptotic case and two decoy states with statistical fluctuation $\mu=0.77$, and for Wang's asymptotic case $\mu=0.43$, which are optimized.

In summary, our investigation demonstrates that a simple two-decoy-state protocol with vacuum+weak decoy states is highly practical and can achieve unconditional security for long-distance (over 100 km) QKD, even with only a few hours of experimental data.

As a final note, we have also studied a simple one-decoy-state protocol. Recently, we have experimentally implemented our one-decoy-state protocol over 15 km of Telecom fibers [15], thus demonstrating the feasibility of our proposal.

ACKNOWLEDGMENTS

This work was financially supported in part by Canadian NSERC, Canada Research Chairs Program, Connaught Fund, Canadian Foundation for Innovation, Ontario Innovation Trust, Premier's Research Excellence Award, Canadian Institute for Photonics Innovations, and University of Toronto startup grant. We acknowledge enlightening discussions with many colleagues including, for example, Charles Bennett, Jean Christian Boileau, Gilles Brassard, Kai Chen, Frédéric Dupuis, Daniel Gottesman, Jim Harrington, Won-Young Hwang, Daniel Lidar, Jeff Kimble, Stephen M. S. Lee, Debbie Leung, Norbert Lütkenhaus, John Preskill, Aephraim Steinberg, Kiyoshi Tamaki, Xiang-Bin Wang, and Zhiliang Yuan. H.-K.L. also acknowledges travel support from the Isaac Newton Institute, Cambridge, U.K., for its quantum information program and from the Institute for Quantum Information at the California Institute of Technology through the National Science Foundation under Grant No. EIA-0086038.

APPENDIX

In this appendix, we will prove that the vacuum+weak decoy protocol is optimal among the two-weak-decoy protocols. We do so by proving that, for a fixed ν_1 (which is larger than ν_2), the lower bound Y_1^{L,ν_1,ν_2} can be no greater than $Y_1^{L,\nu_1,0}$ [see Eq. (A8), and the upper bound $e_1^{U,\nu 1,\nu 2}$ can be no less than $e_1^{U,\nu 1,0}$ [see Eq. (A10)]. We will consider those bounds as given in Eqs. (21) and (25). In what follows, we assume the conditions given by Eq. (15),

$$0 \leqslant \nu_2 < \nu_1,$$

$$\nu_1 + \nu_2 < \mu. \tag{A1}$$

Theorem. Given μ, ν_1, η, Y_0, and $e_{detector}$, the lower bound of Y_1 given in Eq. (21),

$$Y_1^{L,\nu_1,\nu_2} = \frac{\mu}{\mu\nu_1 - \mu\nu_2 - \nu_1^2 + \nu_2^2} \times \left(Q_{\nu_1} e^{\nu_1} - Q_{\nu_2} e^{\nu_2} - \frac{\nu_1^2 - \nu_2^2}{\mu^2} Q_\mu e^\mu \right),$$

is a *decreasing* function of ν_2, and the upper bound of e_1 given in Eq. (25),

$$e_1^{U,\nu_1,\nu_2} = \frac{E_{\nu_1} Q_{\nu_1} e^{\nu_1} - E_{\nu_2} Q_{\nu_2} e^{\nu_2}}{(\nu_1 - \nu_2) Y_1^{L,\nu_1,\nu_2}},$$

is an *increasing* function of ν_2, under conditions Eq. (48). Here Q_μ, Q_{ν_1}, Q_{ν_2}, E_μ, E_{ν_1}, and E_{ν_2} are given by Eqs. (10) and (11).

Proof. First we will prove that Y_1^{L,ν_1,ν_2} is a decreasing function of ν_2 and then prove that e_1^{U,ν_1,ν_2} is an increasing function of ν_2.

Define functions $G(\mu)$ and $J(\mu)$ as

$$G(\mu) = Q_\mu e^\mu = (Y_0 + 1 - e^{-\eta\mu}) e^\mu,$$

$$J(\mu) = E_\mu Q_\mu e^\mu = [e_0 Y_0 + e_{detector}(1 - e^{-\eta\mu})] e^\mu.$$

Take the first derivative of $G(\mu)$ and $J(\mu)$,

$$G'(\mu) = Q_\mu e^\mu + \eta e^{(1-\eta)\mu},$$

$$J'(\mu) = E_\mu Q_\mu e^\mu + \eta e_{detector} e^{(1-\eta)\mu},$$

which are both increasing functions, and $G'(\mu) \geqslant 0$, $J'(\mu) \geqslant 0$. By mathematical induction, it is not difficult to prove the following claim.

Claim 1. For any order derivative of $G(\mu)$ and $J(\mu)$, $G^{(n)}(\mu) \geqslant 0$ and $J^{(n)}(\mu) \geqslant 0$ are increasing functions.

Some useful inequalities. With Claim 1 and the Taylor Series of $G(\mu)$, we have

$$G(\mu) = \sum_{i=0}^{i=\infty} G^{(i)}(\mu) \frac{\mu^i}{i!} \geqslant \mu G'(\mu). \tag{A2}$$

According to the mean value theorem,

$$\frac{G(\nu_1)-G(\nu_2)}{\nu_1-\nu_2}=G'(\nu_3),$$

$$\frac{J(\nu_1)-J(\nu_2)}{\nu_1-\nu_2}=J'(\nu_4), \tag{A3}$$

where $\nu_3,\nu_4\in[\nu_2,\nu_1]$. Because $G'(\mu)$ and $J'(\mu)$ are increasing functions, we can bound Eq. (A3),

$$G'(\nu_2)\leqslant\frac{G(\nu_1)-G(\nu_2)}{\nu_1-\nu_2}\leqslant G'(\nu_1), \tag{A4}$$

$$J'(\nu_2)\leqslant\frac{J(\nu_1)-J(\nu_2)}{\nu_1-\nu_2}\leqslant J'(\nu_1). \tag{A5}$$

Similarly,

$$G''(\nu_2)\leqslant\frac{G'(\nu_1)-G'(\nu_2)}{\nu_1-\nu_2}\leqslant G''(\nu_1). \tag{A6}$$

Define a function

$$F(\nu_2)=\frac{1}{\mu-\nu_1-\nu_2}\left(Q_\mu e^\mu-\frac{\mu}{\nu_1-\nu_2}(Q_{\nu_1}e^{\nu_1}-Q_{\nu_2}e^{\nu_2})\right)$$
$$=\frac{1}{\mu-\nu_1-\nu_2}\left(G(\mu)-\frac{\mu}{\nu_1-\nu_2}[G(\nu_1)-G(\nu_2)]\right).$$

Claim 2. The function $F(\nu_2)$ is an increasing function of ν_2, under the conditions given in Eq. (A1).

Proof. To determine if the function is increasing or decreasing we will need the derivative:

$$F'(\nu_2)=\frac{1}{(\mu-\nu_1-\nu_2)^2}\left(G(\mu)-\frac{\mu}{\nu_1-\nu_2}[G(\nu_1)-G(\nu_2)]\right)-\frac{1}{\mu-\nu_1-\nu_2}\frac{\mu}{(\nu_1-\nu_2)^2}[G(\nu_1)-G(\nu_2)]+\frac{1}{\mu-\nu_1-\nu_2}\frac{\mu}{\nu_1-\nu_2}G'(\nu_2)$$
$$\geqslant\frac{1}{(\mu-\nu_1-\nu_2)^2}[G(\mu)-\mu G'(\nu_1)]-\frac{1}{\mu-\nu_1-\nu_2}\frac{\mu}{\nu_1-\nu_2}G'(\nu_1)+\frac{1}{\mu-\nu_1-\nu_2}\frac{\mu}{\nu_1-\nu_2}G'(\nu_2)$$
$$\geqslant\frac{1}{(\mu-\nu_1-\nu_2)^2}[\mu G'(\mu)-\mu G'(\nu_1+\nu_2)]-\frac{\mu}{\mu-\nu_1-\nu_2}G''(\nu_1)\geqslant\frac{\mu}{\mu-\nu_1-\nu_2}[G''(\nu_1+\nu_2)-G''(\nu_1)]\geqslant 0. \tag{A7}$$

Here, to prove the first inequality, we have made use of Eq. (A4); to prove the second inequality, we have made use of Eqs. (A2) and (A6) and Claim 1; to prove the third inequality, we have made use of Eq. (A6); to prove the last inequality, we have made use of Claim 1.

Proof that Y_1^{L,ν_1,ν_2} is a decreasing function. Rewrite the lower bound of Y_1, in Eq. (21),

$$Y_1^{L,\nu_1,\nu_2}=\frac{\mu}{\mu\nu_1-\mu\nu_2-\nu_1^2+\nu_2^2}\left(Q_{\nu_1}e^{\nu_1}-Q_{\nu_2}e^{\nu_2}-\frac{\nu_1^2-\nu_2^2}{\mu^2}Q_\mu e^\mu\right)$$
$$=\frac{\mu}{\mu\nu_1-\mu\nu_2-\nu_1^2+\nu_2^2}(Q_{\nu_1}e^{\nu_1}-Q_{\nu_2}e^{\nu_2})-\frac{\mu}{\mu\nu_1-\mu\nu_2-\nu_1^2+\nu_2^2}\frac{\nu_1^2-\nu_2^2}{\mu^2}Q_\mu e^\mu=\frac{\mu}{\mu-\nu_1-\nu_2}\frac{Q_{\nu_1}e^{\nu_1}-Q_{\nu_2}e^{\nu_2}}{\nu_1-\nu_2}$$
$$-\frac{\nu_1+\nu_2}{\mu-\nu_1-\nu_2}\frac{Q_\mu e^\mu}{\mu}$$
$$=\frac{\mu}{\mu-\nu_1-\nu_2}\frac{Q_{\nu_1}e^{\nu_1}-Q_{\nu_2}e^{\nu_2}}{\nu_1-\nu_2}-\left(\frac{1}{\mu-\nu_1-\nu_2}-\frac{1}{\mu}\right)Q_\mu e^\mu$$
$$=\frac{1}{\mu}Q_\mu e^\mu-\frac{1}{\mu-\nu_1-\nu_2}\left(Q_\mu e^\mu-\frac{\mu}{\nu_1-\nu_2}(Q_{\nu_1}e^{\nu_1}-Q_{\nu_2}e^{\nu_2})\right)=\frac{1}{\mu}Q_\mu e^\mu-F(\nu_2). \tag{A8}$$

With Claim 2, we show that Y_1^{L,ν_1,ν_2} is a decreasing function of ν_2.

Define a function

$$K(\nu_2)=\frac{E_{\nu_1}Q_{\nu_1}e^{\nu_1}-E_{\nu_2}Q_{\nu_2}e^{\nu_2}}{\nu_1-\nu_2}=\frac{J(\nu_1)-J(\nu_2)}{\nu_1-\nu_2}.$$

Claim 3. The function $K(\nu_2)$ is an increasing function with ν_2.

Proof. To determine if the function is increasing or de-creasing we will need the derivative:

$$K'(\nu_2) = \frac{J(\nu_1) - J(\nu_2)}{(\nu_1 - \nu_2)^2} - \frac{J'(\nu_2)}{\nu_1 - \nu_2} \geqslant \frac{J'(\nu_2)}{\nu_1 - \nu_2} - \frac{J'(\nu_2)}{\nu_1 - \nu_2} = 0, \tag{A9}$$

where the first inequality is due to Eq. (A5).

Proof that e_1^{U,ν_1,ν_2} is an increasing function. Reform the lower bound of e_1, in Eq. (25),

$$e_1^{U,\nu_1,\nu_2} = \frac{E_{\nu_1} Q_{\nu_1} e^{\nu_1} - E_{\nu_2} Q_{\nu_2} e^{\nu_2}}{(\nu_1 - \nu_2) Y_1^{L,\nu_1,\nu_2}} = \frac{K(\nu_2)}{Y_1^{L,\nu_1,\nu_2}}. \tag{A10}$$

With Claim 3 and the decreasing function of Y_1^{L,ν_1,ν_2}, we show that e_1^{U,ν_1,ν_2} is an increasing function of ν_2.

In summary, we have proved the theorem.

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

[2] D. Mayers, J. ACM **48**, 351 (2001); preliminary version in *Advances in Cryptology—Proceedings of Crypto '96*, edited by N. Koblitz, Lecture Notes in Computer Science Vol. 1109 (Springer-Verlag, New York, 1996), pp. 343–357;H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999);E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury, in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC'00)* (ACM Press, New York, 2000), pp. 715–724;M. Ben-Or, available at http://www.msri.org/publications/ln/msri/2002/qip/ben-or/1/

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

[4] A. K. Ekert and B. Huttner, J. Mod. Opt. **41**, 2455 (1994); D. Deutsch *et al.*, Phys. Rev. Lett. **77**, 2818 (1996); **80**, 2022(E) (1998).

[5] C. Gobby, Z. L. Yuan, and A. J. Shields, Appl. Phys. Lett. **84**, 3762 (2004).

[6] Kimura, T. *et al.*, Jpn. J. Appl. Phys., Part 2 **43**, 1217 (2004).

[7] D. Gottesman, H.-K. Lo, Norbert Lutkenhaus, and John Preskill, Quantum Inf. Comput. **4**, 325 (2004).

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

[9] H.-K. Lo, X. Ma, and K. Chen, Phys. Rev. Lett. **94**, 230504 (2005). Preliminary results were presented in *Proceedings of IEEE ISIT 2004* (IEEE Press, Chicago, 2004), p. 137; and at http://www.fields.utoronto.ca/programs/scientific/04-05/quantumIC/abstracts/lo.ppt;see also X. Ma, e-print quant-ph/0503057.

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

[11] Here, the gain of a particular type of signal is defined to be the fraction of detection events by Bob that is due to that particular type of signal.

[12] G. Brassard and L. Salvail, in *Advances in Cryptology EUROCRYPT '93*, edited by T. Helleseth, Lecture Notes in Computer Science Vol. 765 (Springer, Berlin, 1994), pp. 410–423.

[13] X. -B. Wang, Phys. Rev. Lett. **94**, 230503 (2005).; e-print quant-ph/0411047.

[14] J. W. Harrington, J. M. Ettinger, R. J. Hughes, and J. E. Nordholt, e-print quant-ph/0503002.

[15] Y. Zhao, B. Qi, X. Ma, H.-K. Lo, and L. Qian, e-print quant-ph/0503192.

[16] C.H. Bennett, Phys. Rev. Lett. **68**, 3121(1992).

[17] M. Koashi, e-print quant-ph/0403131.

[18] Norbert Lütkenhaus, Phys. Rev. A **61**, 052304 (2000).

[19] M. Bourennane, F. Gibson, A. Karlsson, A. Hening, P. Jonsson, T. Tsegaye, D. Ljunggren, and E. Sundberg, Opt. Express **4**, 383 (1999).

[20] Zhiliang Yuan (private communication).