Simulating single photons with realistic photon sources

Xiao Yuan,[1] Zhen Zhang,[1] Norbert Lütkenhaus,[2] and Xiongfeng Ma[1]
[1]*Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, China*
[2]*Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario, Canada N2L3G1*
(Received 6 September 2016; published 6 December 2016)

Quantum information processing provides remarkable advantages over its classical counterpart. Quantum optical systems have been proved to be sufficient for realizing general quantum tasks, which, however, often rely on single-photon sources. In practice, imperfect single-photon sources, such as a weak-coherent-state source, are used instead, which will inevitably limit the power in demonstrating quantum effects. For instance, with imperfect photon sources, the key rate of the Bennett-Brassard 1984 (BB84) quantum key distribution protocol will be very low, which fortunately can be resolved by utilizing the decoy-state method. As a generalization, we investigate an efficient way to simulate single photons with imperfect ones to an arbitrary desired accuracy when the number of photonic inputs is small. Based on this simulator, we can thus replace the tasks that involve only a few single-photon inputs with the ones that make use of only imperfect photon sources. In addition, our method also provides a quantum simulator to quantum computation based on quantum optics. In the main context, we take a phase-randomized coherent state as an example for analysis. A general photon source applies similarly and may provide some further advantages for certain tasks.

DOI: 10.1103/PhysRevA.94.062305

I. INTRODUCTION

Quantum information science has developed rapidly in the last few decades. At the theoretical level, various schemes have been proposed to solve classical intractable problems or provide certain quantum advantages. Specifically, Shor's factorization algorithm [1] indicates that quantum computing can exponentially enhance the computational power in certain tasks compared to a classical computer. In addition, quantum key distribution (QKD) protocols [2,3] enable remote users to extend secret keys with security guaranteed by the basic principles of quantum mechanics.

In experiment, quantum optics is favored for realizing quantum information processing tasks due to the weak interaction between a photon and its environment. Especially in quantum communication, various tasks, such as long-distance quantum key distribution [4] and quantum teleportation [5], are realized with linear optics. On the other hand, linear optics is not enough to realize universal quantum computation. Roughly speaking, it requires exponentially large resources to implement a linear quantum optical computer [6]. Thus, nonlinearity is crucial for universal quantum computation in linear optics. One possible way is to use nonlinear optics [7], which still faces the scalable difficulty for current technology. On the other hand, Knill, Laflamme, and Milburn (KLM) [8] have shown that efficient quantum computation is possible using only linear optics with single-photon sources. The nonlinearity is introduced by adaptive measurements which can be realized with the techniques of quantum teleportation [9].

In reality, perfect single-photon sources and detectors are not available. Instead, other imperfect photon sources, such as heralded spontaneous parametric down-conversion (SPDC) sources, are used to simulate single photons. Meanwhile, single-photon detectors generally have low efficiency. The imperfection of devices will lead to unexpected events, thus limiting the quantum advantage. In experiment, entangling eight photons is the best reported result [10,11].

The imperfections of single-photon sources not only affect the accuracy but cause loopholes in cryptography protocols. Specifically, the multiphoton parts will lead to photon-number-splitting attacks [12], which makes the key rate of the well-known Bennett-Brassard 1984 (BB84) QKD protocol [2] very low. The imperfect photon sources seem to limit the power of optical realization of quantum information processing. Surprisingly, this is not the case in reality. Even with imperfect photon sources, such as a weak coherent state, as an input, secure QKD protocols are still possible by utilizing the decoy-state method [13–15]. By inputting two or more coherent states, one can still estimate the information leaked in eavesdropping and thus make the whole process secure.

In this work, we generalize the idea of a decoy state to general optical circuits. As an example, we show that it is possible to simulate a single photon to arbitrary accuracy efficiently by making efficient use of phase-randomized coherent states. In addition, we generalize our result to multiple photons. We show that replacing a few single photons with coherent states is possible in general quantum information tasks. For large numbers of photons, we link our work to the scenario of quantum computation. Our method thus provides a quantum simulator to general quantum computing processes. Finally, we discuss that our method works for general photon sources.

II. FRAMEWORK

In this section, we first review the basic framework of optical circuits. With a single photon as the input, whose density matrix is denoted by ρ_{in}, a general optical circuit can be regarded as a quantum channel described in Fig. 1, which involves a unitary interaction U between the signal photon and the environment E. After the channel, a measurement M is performed on the output photon ρ_{out}.

Such a general quantum channel can be fitted into many scenarios such as QKD, where Alice encodes her information in the input signal state ρ_{in} and sends it through a public quantum channel to Bob. On the output side, Bob performs photon-number measurement on his received signal ρ_{out},

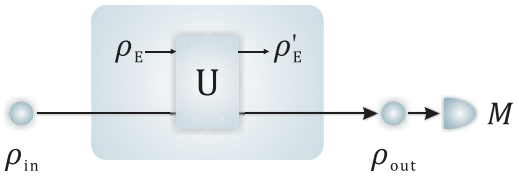

FIG. 1. Optical circuits with a single photon as input.

denoted by a positive operator-valued measure (POVM). For a specific POVM element M, the detection probability is

$$Q = \mathrm{Tr}[M\rho_{\mathrm{out}}], \tag{1}$$

where $\rho_{\mathrm{out}} = \mathrm{Tr}_E[U(\rho_{\mathrm{in}} \otimes \rho_{\mathrm{E}})U^\dagger]$. Necessary classical post-processing should be applied to the outcome Q to extract the final desired quantum information. In the following, we will focus on this specific POVM element; it is straightforward to see that our results apply similarly to the other POVM elements.

Ideally, the information is encoded on a single photon. However, a more general scenario is that Alice feeds a mixture of Fock states,

$$\rho'_{\mathrm{in}} = \sum_{k=0}^{\infty} P(k)|k\rangle\langle k|, \tag{2}$$

where $|k\rangle$ represents a Fock state that contains k photons and $P(k)$ is the photon-number distribution satisfying $P(k) \in [0,1]$ and $\sum_{k=0}^{\infty} P(k) = 1$. For a single-photon source, $P(k)$ is a Kronecker δ function,

$$P(k) = \delta_{k1} = \begin{cases} 1, & k = 1, \\ 0, & k \neq 1. \end{cases} \tag{3}$$

For a phase-randomized coherent-state source, $P(k)$ is a Poisson distribution [14],

$$P(k) = \frac{\mu^k}{k!} e^{-\mu}. \tag{4}$$

When the general photon state defined in Eq. (2) is used, Bob's detection probability defined in Eq. (1) is given by

$$Q = \sum_{k=0}^{\infty} P(k)P(\mathrm{click}|k\ \mathrm{photons}), \tag{5}$$

where $P(\mathrm{click}|k\ \mathrm{photons})$ denotes the detection probability with k photons as input, that is,

$$P(\mathrm{click}|k\ \mathrm{photons}) = \mathrm{Tr}[MU(|k\rangle\langle k| \otimes \rho_{\mathrm{E}})U^\dagger]. \tag{6}$$

For easier presentation, we will denote $P(\mathrm{click}|k\ \mathrm{photons})$ by Y_k. When $k = 0$, Y_0 corresponds to the yield with no photon input, i.e., the dark count. When $k = 1$, Y_1 denotes the probability with a single photon as input. For a single-photon source, we have $Q = Y_1$. However, for a phase-randomized coherent source, we have

$$Q = \sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} Y_k. \tag{7}$$

Ideally, a single-photon source is required for several quantum information processing tasks. The accurate probability distribution Y_1 can be obtained only with a single-photon source. However, if we intend to learn only the value of Y_k, we show in this work that such a value can be accurately and efficiently estimated with several phase-randomized coherent states as input.

III. SIMULATING A SINGLE PHOTON

In this section, we will focus on simulating the probability distribution Y_1 with a single photon of a general quantum circuit. For easier presentation, let us first define $A_\mu = Qe^\mu$, that is,

$$A_\mu = \sum_{k=0}^{\infty} \frac{\mu^k}{k!} Y_k. \tag{8}$$

To estimate Y_1, we will make use of the idea of the decoy-state method originally applied in QKD [16]. That is, Alice chooses a few probe intensities of phase-randomized coherent states to get several detections A_μ. By regarding the probability Y_i with i photons for $i = 1,2,\ldots$ as an unknown variable, we thus get several linear equations of Y_i in the form of Eq. (8) with different μ and A_μ. As there is an infinite number of unknown variables, we need an infinite number of equations to deterministically decide Y_1. However, we can still approximately estimate Y_1 with finite linear equations. With a greater number of coherent states used, the estimation becomes more accurate. A similar analysis has been done for QKD with a few decoy states [17]; however, the obtained estimation is not optimal.

For heuristic presentation, we will show examples of estimating Y_1 with one and two probe intensities. Another example with three probe intensities can be found in Appendix. A. Furthermore, we analytically derive an estimator for Y_1 with L general probe intensities plus one vacuum intensity. With an explicit example, we numerically prove that the estimation error decays exponentially with the number of probes L. These results are derived in the asymptotic case, where for each probe intensity there is an infinite number of samples. To take the finite-size effect into account, we consider a finite number of samples $M/(L+1)$ for each probe intensity. By considering the total error of our method, including estimation and statistical errors, we find that it scales inversely proportionally to a power function of the total number of coherent pulses M, $\exp[-O(\ln(M))]$. Therefore, our method with coherent probes is efficient.

In the following, we will first discuss the case with one, two, and three probe intensities and then generalize the result to L probe intensities. The discussion of one, two, and three probe intensities can be found in the decoy-state method in QKD [16], and generalization to L probe intensities is our result.

A. Review: One probe intensity

First, we consider that only one phase-randomized coherent state ρ_μ is used. In this case, an estimation of Y_1 is given by the redefined probability A_μ in Eq. (8) divided by its intensity μ, that is, $Y_1^{\mathrm{est}} = A_\mu/\mu$. To see the estimation accuracy, we

use the relation between A_μ and Y_1 via Eq. (8),

$$Y_1 = \frac{1}{\mu}\left(A_\mu - Y_0 - \frac{\mu^2}{2}Y_2 - \cdots\right) = Y_1^{\text{est}} - \frac{1}{\mu}\left(Y_0 + \frac{\mu^2}{2}Y_2 + \cdots\right). \tag{9}$$

As Y_n corresponds to the probability with n photons as input, we have $Y_n \in [0,1]$ and hence the bounds of Y_1,

$$Y_1^{\text{est}} - \frac{e^\mu - \mu}{\mu} \leqslant Y_1 \leqslant Y_1^{\text{est}}. \tag{10}$$

The estimation accuracy is defined by the interval between the upper and lower bounds,

$$\Delta_0 = \frac{e^\mu - \mu}{\mu}, \tag{11}$$

which is minimized at $\mu = 1$ with the value of $e - 1 > 1$. Thus, at least one of the bounds in Eq. (10) is trivial since $Y_1 \in [0,1]$.

It is easy to see that a single use of a coherent state gives a very loose estimation of Y_1. This can be intuitively understood by the dark-count contribution Y_0 in Eq. (9). To overcome this, we can input an additional coherent state and show in the following that Y_1 can be estimated to much better accuracy.

B. Review: Vacuum + one probe intensities

Now, suppose Alice can add another probe coherent state ρ_ν. In this scenario, there are two linear equations,

$$A_\mu = Y_0 + \mu Y_1 + \frac{\mu^2}{2}Y_2 + \cdots, \quad A_\nu = Y_0 + \nu Y_1 + \frac{\nu^2}{2}Y_2 + \cdots. \tag{12}$$

Subtracting one from the other, we have

$$\frac{A_\mu - A_\nu}{\mu - \nu} = Y_1 + \frac{\mu + \nu}{2}Y_2 + \cdots. \tag{13}$$

Therefore, we can estimate Y_1 by $Y_1^{\text{est}} = \frac{A_\mu - A_\nu}{\mu - \nu}$ and have the relation

$$Y_1 = Y_1^{\text{est}} - \frac{\mu + \nu}{2}Y_2 - \cdots. \tag{14}$$

Assuming $\mu > \nu$, the bounds of Y_1 are given by

$$Y_1^{\text{est}} - \frac{e^\mu - \mu - e^\nu + \nu}{\mu - \nu} \leqslant Y_1 \leqslant Y_1^{\text{est}}. \tag{15}$$

The size of the interval is

$$\Delta_1 = \frac{e^\mu - e^\nu}{\mu - \nu} - 1. \tag{16}$$

The minimum of Δ_1 is reached for $\nu = 0$, and it increases with μ. For a small μ, we can approximate the interval by

$$\Delta_1 = \frac{\mu}{2} + O\left(\frac{\mu^2}{2!}\right). \tag{17}$$

The intuition behind the choices of the intensities comes from the motivation to estimate the background contribution Y_0. After that, the estimation error of Y_1 suffers only from contributions of more than two photon numbers, that is, $O(\mu^2/2!)$. Therefore, in the following, we will always consider the vacuum probe intensity.

C. Vacuum + L probe intensities

We leave the result with the vacuum + two probe intensities to Appendix A and consider a general case where Alice inputs an $L+1$ phase-randomized coherent state ρ_{μ_0} (vacuum), $\rho_{\mu_1}, \ldots, \rho_{\mu_L}$. Suppose $\mu_0 = 0$ and $\mu_1 < \mu_2 < \cdots < \mu_L$; an estimation of Y_1 is given in Appendix B [Eq. (B11)] by

$$Y_1^{\text{est}} = \mu_1\mu_2\cdots\mu_L \sum_{j=1}^{L} \frac{\mu_j^{-2}\left(A_{\mu_j} - A_0\right)}{\prod_{1\leqslant n\leqslant L; n\neq j}(\mu_n - \mu_j)}, \tag{18}$$

where A_{μ_j} is the gain of the coherent-state input with intensity μ_j. The bounds of the Y_1 estimation are

$$Y_Y^{\text{est}} - \Delta_L \leqslant Y_1 \leqslant Y_1^{\text{est}}, \quad Y_1^{\text{est}} \leqslant Y_1 \leqslant Y_1^{\text{est}} + \Delta_L \tag{19}$$

for L to be odd and even, respectively, where the interval between the upper and lower bounds is given according to Eq. (B16) by

$$\Delta_L = (-1)^{L+1}\left(\mu_1\mu_2\cdots\mu_L \sum_{j=1}^{L} \frac{\mu_j^{-2}(e^{\mu_j} - 1)}{\prod_{1\leqslant n\leqslant L; n\neq j}(\mu_n - \mu_j)} - 1\right) = \frac{\mu_1\cdots\mu_L}{(L+1)!} + O\left[\frac{\mu_1\cdots\mu_L\sum\mu_l}{(L+2)!}\right]. \tag{20}$$

When the intensities μ_j are small, we can see that the estimation interval exponentially decreases with L. Thus, a single photon can be efficiently simulated with a coherent source as the input.

According to Eqs. (B17) and (B19), the estimation Y_1^{est} and the interval Δ_L can be represented as a linear combination of A_{μ_j} as

$$Y_1^{\text{est}} = \sum_{j=1}^{\lceil L/2\rceil} \lambda_{2j-1}A_{\mu_{2j-1}} - \sum_{j=1}^{\lfloor L/2\rfloor} \lambda_{2j}A_{\mu_{2j}} + \lambda_0 A_0,$$
$$\Delta_L = (-1)^{L+1}\left(\sum_{j=1}^{\lceil L/2\rceil} \lambda_{2j-1}e^{\mu_{2j-1}} - \sum_{j=1}^{\lfloor L/2\rfloor} \lambda_{2j}e^{\mu_{2j}} + \lambda_0 - 1\right), \tag{21}$$

where the coefficients λ_j are positive and given by

$$\lambda_0 = \sum_{j=1}^{L}(-1)^j\lambda_j, \quad \lambda_j = \frac{(-1)^{j+1}}{\mu_j}\prod_{1\leqslant n\leqslant L; n\neq j}\frac{\mu_n}{(\mu_n - \mu_j)}, \quad 1 \leqslant j \leqslant L. \tag{22}$$

We refer to Appendix B for the derivation of the results and focus on the performance.

D. Total error of estimation

In the estimation of Y_1 given in Eq. (18), we assume A_{μ_j} to be accurate. In practice, we have to input several

copies of the same coherent state with intensity μ_j, and A_{μ_j} can be estimated from the measurement. In this case, besides the estimation error Δ_L, we have to consider the statistical error of estimating each A_{μ_j}. In the last part, we proved that the estimation with a few probes intensities can efficiently simulate the result with a single-photon state in the asymptotical scenario. In the following, we will show that such a method is also efficient when focusing on finite data size.

To show the method is as efficient as the with one with a single photon, we consider independent and identically distributed (i.i.d.) sampling for simplicity. In QKD, such finite-size effects without assumptions have been analyzed for the vacuum plus weak-decoy-state formalism [18–20]. In Ref. [20], it was shown that the difference is only a factor when the sample size is large. Thus, we leave the analysis without additional assumptions to future work.

Under the i.i.d. assumption, the statistical error $\Delta_s(A)$ of A_μ can be approximated by

$$\Delta_s(A) \lesssim \frac{1}{\sqrt{m}} = \sqrt{\frac{L+1}{M}}, \tag{23}$$

where m is the number of samples for each μ_j and $M = m(L+1)$ is the total number of samples. Here, we consider the same statistical error estimation for all A_μ for simplicity. A tighter bound that involves A_μ can be further applied when the value of A_μ is known.

The sampling-induced error of Y_1^{est} is given by

$$\Delta_s\left(Y_1^{\text{est}}\right) = \Delta_s(A)\sqrt{\sum_{j=0}^{L}\lambda_j^2}. \tag{24}$$

The total error in experiment is thus

$$\Delta_t \lesssim \Delta_s + \Delta_L. \tag{25}$$

In the following, we will give an example to show the scale of total error Δ_t compared to a fixed total sample size M. With a perfect single-photon source, the total error scales as $O(1/\sqrt{M})$. With a phase-randomized coherent state, we show that by inputting appropriate quiz states, the total error also scales as a power function of M.

E. Example

As different probe intensities will lead to different estimation errors Δ_L, we take only an example with probe intensities $\mu_j = j/L$ for $j = 0,1,\ldots,L$. There may exist better choices of the probe intensities that cause a smaller total error. In our example, the coefficients λ_j, defined in Eq. (22), can be calculated by

$$\begin{aligned}\lambda_j &= \frac{(-1)^{j+1}}{j/L}\prod_{1\leqslant n\leqslant L; n\neq j}\frac{n/L}{(n/L - j/L)}\\ &= \frac{L(-1)^{j+1}}{j}\prod_{1\leqslant n\leqslant L; n\neq j}\frac{n}{(n-j)}\\ &= \frac{L}{j}\binom{L}{j},\end{aligned} \tag{26}$$

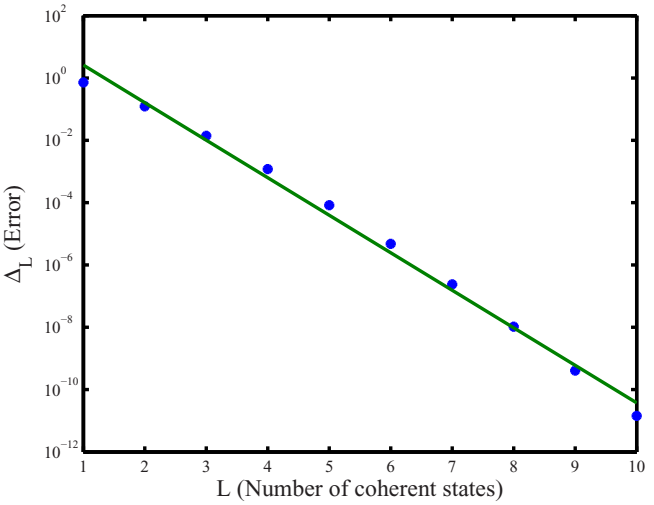

FIG. 2. Error Δ_L for specific usage of coherent intensities, $\mu_j = j/L$ for $j = 0,1,\ldots,L$. L runs only from 1 to 10 for computational accuracy limit. Blue dots are the Δ_L value, and the green line is the exponential fit.

and λ_0 can be calculated by

$$\lambda_0 = \sum_{j=1}^{L}(-1)^j\frac{L}{j}\binom{L}{j}. \tag{27}$$

In this case, the estimation error Δ_L can be numerically calculated as shown in Fig. 2. A linear fitting between $\ln\Delta_L$ and L thus gives the relation $\ln\Delta_L = -2.772L + 3.718$. It is straightforward to see that with increasing L, the interval Δ_L exponentially approaches zero.

The sampling error is

$$\Delta_s\left(Y_1^{\text{est}}\right) = \Delta_s(A)f(L), \tag{28}$$

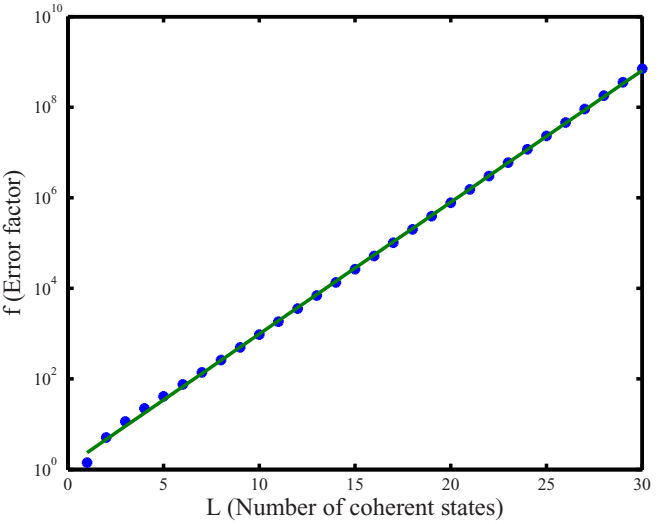

FIG. 3. Error factor f for different L. Blue dots are the Δ_L value, and the green line is the exponential fit.

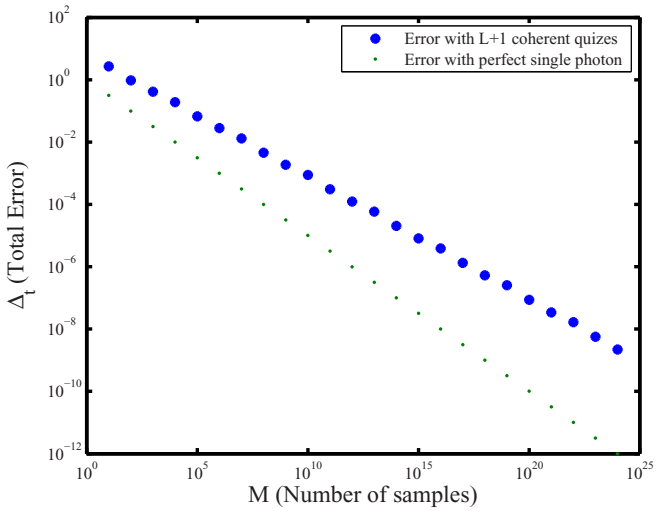

FIG. 4. Optimized total error for different numbers of samples. The blue dots are the total error for $L(M)$ coherent probe intensities. The green dots are the total error for a single-photon source, that is, $1/\sqrt{M}$.

where $f(L)$ is a constant factor

$$f(L) = \sqrt{\left[\sum_{j=1}^{L}(-1)^j \frac{L}{j}\binom{L}{j}\right]^2 + \sum_{j=1}^{L}\left[\frac{L}{j}\binom{L}{j}\right]^2}. \quad (29)$$

As shown in Fig. 3, the error factor f is roughly exponential to the number of coherent states L. A linear fitting thus gives the relation $\ln f = 0.67L + 0.189$.

Thus, the total error can be approximately modeled by

$$\Delta_t = \sqrt{\frac{L+1}{M}} e^{0.67L+0.189} + e^{-2.772L+3.718}. \quad (30)$$

For a given total number of samples M, we can optimize over L to minimize the total error Δ_t. We solve this problem numerically. As shown in Fig. 4, the total error Δ_t is still inversely proportional to a power function of the number of samples M. That is, we have $\Delta_t \approx 6.6128/M^{0.3931}$, which fits the data we present in Fig. 4.

In addition, the optimized number of probe intensities is shown in Fig. 5. Roughly speaking, L is linearly proportional to $\ln M$, which explains why Δ_t is still a power function of M.

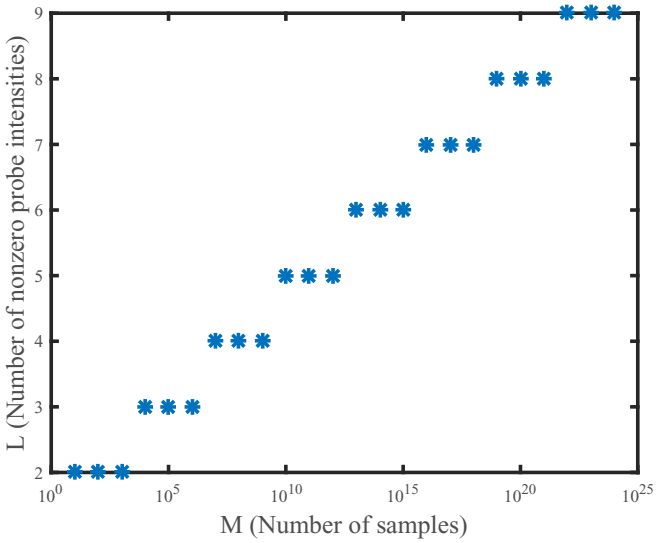

FIG. 5. Optimal number of nonzero probe intensities L for different numbers of samples M.

IV. PARAMETER ESTIMATION: MULTIPLE INPUT MODES

Now, we consider a general optical circuit with n distinguishable photons as inputs. The quantum circuits can be well described by a quantum channel with n optical modes, as shown in Fig. 6. The input state ρ_{in} consists of n single photons which correspond to each of the input modes. After a unitary interaction between the input particles and the environment described by ρ_E, measurements are performed on each of the output modes. For each of the measurements M_i, with $i = 1,2,\ldots,n$, the detection probability is given by

$$Q_i = \text{Tr}_{\neq i,E}[M_i \rho_{\text{out}}] = \text{Tr}_{\neq i,E}[M_i U(\rho_{\text{in}} \otimes \rho_{\text{E}})U^\dagger], \quad (31)$$

where the trace is over the environment E and all the input modes except the ith one.

In the previous section, we showed that each single-mode photon can be simulated efficiently with multiple usages of coherent pulses. Here, we generalize the result to the n-input-mode case. We define a coincidence detection by

$$Q = \text{Tr}[M\rho_{\text{out}}] = \text{Tr}[MU(\rho_{\text{in}} \otimes \rho_{\text{E}})U^\dagger], \quad (32)$$

where the measurement is $M = M_1 \otimes M_2 \otimes \cdots \otimes M_n$. When the input state is a mixture of photon-number states,

$$\rho_{\text{in}} = \sum_{k_1,k_2,\ldots,k_n=0}^{\infty} P(k_1,k_2,\ldots,k_n)|k_1k_2\cdots k_n\rangle\langle k_1k_2\cdots k_n|, \quad (33)$$

the coincidence detection can be expressed by

$$Q = \sum_{k_1,k_2,\ldots,k_n=0}^{\infty} P(k_1,k_2,\ldots,k_n)Y_{k_1k_2\cdots k_n}, \quad (34)$$

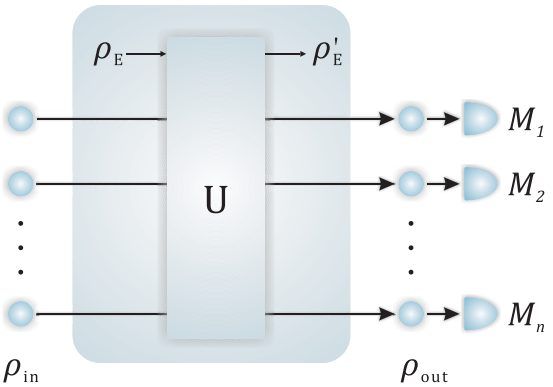

FIG. 6. Schematic for the a general quantum channel with n optical input modes.

where $Y_{k_1k_2\ldots k_n}$ is the coincidence detection probability for the case that the ith mode has k_i photons,

$$Y_{k_1k_2\cdots k_n} = \mathrm{Tr}[MU(|k_1k_2\cdots k_n\rangle\langle k_1k_2\cdots k_n| \otimes \rho_{\mathrm{E}})U^\dagger]. \quad (35)$$

In the following, we show that with coherent state as the input, we can also estimate the coincidence detection for the single-photon input $Y_{11\ldots1}$ to an arbitrary accuracy. For simplicity, we consider the same probe intensities for different input modes. The derivation of different probe intensities for different input modes follows similarly.

A. Two modes with vacuum + one probe intensities

First, we consider only two input optical modes and two probe intensities. From Sec. III B, we find that one of the two probe intensities should be a vacuum state and the other should be a weak state ρ_μ in the optimal case. Like in Eq. (7), we have

$$Q = \sum_{k_1=0}^{\infty} \frac{\mu_1^{k_1}}{k_1!} e^{-\mu_1} \sum_{k_2=0}^{\infty} \frac{\mu_2^{k_2}}{k_2!} e^{-\mu_2} Y_{k_1k_2}, \quad (36)$$

where μ_1 (μ_2) and k_1 (k_2) are the coherent-state intensity and the photon number for the first (second) mode, respectively. Like in Eq. (8), we define $A_{\mu_1\mu_2} = Qe^{\mu_1}e^{\mu_2}$,

$$A_{\mu_1\mu_2} = \sum_{k_1,k_2=0}^{\infty} \frac{\mu_1^{k_1}\mu_2^{k_2}}{k_1!k_2!} Y_{k_1k_2}. \quad (37)$$

When each mode is input with a coherent state with zero and μ intensity, we have four equalities based on Alice's four possible input cases:

$$\begin{aligned}
A_{00} &= Y_{00},\\
A_{\mu 0} &= \sum_{k_1=0}^{\infty} \frac{\mu_1^k}{k_1!} Y_{k_1 0} = Y_{00} + \mu Y_{10} + \frac{\mu^2}{2} Y_{20} + \cdots,\\
A_{0\mu} &= \sum_{k_2=0}^{\infty} \frac{\mu_2^k}{k_2!} Y_{0k_2} = Y_{00} + \mu Y_{01} + \frac{\mu^2}{2} Y_{02} + \cdots, \quad (38)\\
A_{\mu\mu} &= \sum_{k_1,k_2=0}^{\infty} \frac{\mu_1^k \mu^{k'}}{k_1!k_2!} Y_{k_1k_2}\\
&= Y_{00} + \mu Y_{10} + \mu Y_{01} + \mu^2 Y_{11} + \cdots.
\end{aligned}$$

With the attempt to estimate Y_{11}, we can linearly combine Y_{00}, Y_{10}, Y_{01}, and Y_{11} by

$$\begin{aligned}
Y_{11}^{\mathrm{est}} &= A_{\mu\mu} - A_{0\mu} - A_{\mu 0} + A_{00}\\
&= \sum_{k_1,k_2=0}^{\infty} \frac{\mu^{k_1}\mu^{k_2}}{k_1!k_2'!} Y_{k_1k_2'} - \sum_{k_1=0}^{\infty} \frac{\mu^{k_1}}{k_1!} Y_{k_1 0} - \sum_{k_2=0}^{\infty} \frac{\mu^{k_2}}{k_2!} Y_{0k_2} + Y_{00}\\
&= \sum_{k_1,k_2=1}^{\infty} \frac{\mu^{k_1}\mu^{k_2}}{k_1!k_2'!} Y_{k_1k_2'}. \quad (39)
\end{aligned}$$

As the Y are always in $[0,1]$, we can bound Y_{11} by

$$Y_{11}^{\mathrm{est}} - \Delta_{1,2} \leqslant Y_{11} \leqslant Y_{11}^{\mathrm{est}}, \quad (40)$$

with the size of the interval $\Delta_{1,2}$

$$\Delta_{1,2} = \frac{1}{\mu^2} \sum_{n,m=1}^{\infty} \frac{\mu^n\mu^m}{n!m!} - 1 = \frac{(e^\mu - 1)^2}{\mu^2} - 1. \quad (41)$$

Here, the first subscript denotes the number of nonzero probe intensities, and the second subscript denotes the number of input modes. For a small μ, we have

$$\Delta_{1,2} = \mu + O(\mu^2). \quad (42)$$

B. n modes with vacuum + one probe intensities

Next, we generalize the result to the case of n input modes each with two possible (vacuum 0 and weak μ) probe intensities. Denoting the nonzero coherent-state intensity for the ith mode by μ_i, the measurement result is given by, similar to Eq. (36),

$$Q = \sum_{k_1,k_2,\ldots,k_n=0}^{\infty} \frac{\mu_1^{k_1}\mu_2^{k_2}\cdots\mu_n^{k_n}}{k_1!k_2!\cdots k_n!} e^{-(\mu_1+\mu_2+\cdots+\mu_n)} Y_{k_1k_2\cdots k_n}. \quad (43)$$

Like in Eq. (37), we define $A_{\mu\mu\cdots\mu} = Qe^{(\mu_1+\mu_2+\cdots+\mu_n)}$, and we have

$$A_{\mu_1\mu_2\cdots\mu_n} = \sum_{k_1,k_2,\ldots,k_n=0}^{\infty} \frac{\mu_1^{k_1}\mu_2^{k_2}\cdots\mu_n^{k_n}}{k_1!k_2!\cdots k_n!} Y_{k_1k_2\ldots k_n}. \quad (44)$$

For easier presentation, we first introduce an operation on the coincidence probability Y. For $Y_{k_1k_2\cdots k_n}$, we define it as

$$Y_{k_1k_2\cdots k_n} = \bigotimes_{i=1}^{n} Y_{k_i}, \quad (45)$$

where the operation $\bigotimes_{i=1}^n$ denotes subscript combination. The notation Y_{k_i} does not make sense unless the operation $\bigotimes_{i=1}^n$ is applied. With this notation, we can rewrite $A_{\mu_1\mu_2\ldots\mu_n}$ by

$$A_{\mu_1\mu_2\cdots\mu_n} = \bigotimes_{i=1}^{n} A_{\mu_i}, \quad (46)$$

where

$$A_{\mu_i} = \sum_{k_i=0}^{\infty} \frac{\mu_i^{k_i}}{k_i!} Y_{k_i}. \quad (47)$$

Notice that $\bigotimes_{i=1}^n$ can still be regarded as a product operation where the multiplication of Y is replaced by the subscript combination. Thus,

$$\begin{aligned}
\bigotimes_{i=1}^{n} A_{\mu_i} &= \sum_{k_1,k_2,\ldots,k_n=0}^{\infty} \frac{\mu_1^{k_1}\mu_2^{k_2}\cdots\mu_n^{k_n}}{k_1!k_2!\cdots k_n!} \bigotimes_{i=1}^{n} Y_{k_i}\\
&= A_{\mu_1\mu_2\cdots\mu_n}. \quad (48)
\end{aligned}$$

With the same spirit, we can derive

$$\begin{aligned}
\bigotimes_{i=1}^{n} (A_{\mu_i} - A_0) &= \bigotimes_{i=1}^{n} \left(\sum_{k_i=0}^{\infty} \frac{\mu_i^{k_i}}{k_i!} Y_{k_i} - Y_{k_i=0} \right)\\
&= \sum_{k_1,k_2,\ldots,k_n=1}^{\infty} \frac{\mu^{k_1}\mu^{k_2}\cdots\mu^{k_n}}{k_1!k_2!\cdots k_n!} Y_{k_1k_2\cdots k_n}, \quad (49)
\end{aligned}$$

which is a generalization to Eq. (39). Here, we let all μ_i equal the same intensity μ. Like for Eq. (39), it is not hard to see that an estimation of $Y_{11\cdots1}$ is given by

$$Y^{\text{est}}_{11\cdots1} = \bigotimes_{i=1}^{n} \left(A_{\mu_i} - A_0\right). \tag{50}$$

Now, the size of the interval for estimating $Y_{11\cdots1}$ is given by

$$\begin{aligned}\Delta_{1,n} &= \frac{1}{\mu^n} \sum_{k_1,\ldots,k_n=1}^{\infty} \frac{\mu^{k_1}\cdots\mu^{k_n}}{k_1!\cdots k_n!} - 1 \\ &= \frac{(e^{\mu}-1)^n}{\mu^n} - 1 \\ &= \frac{n}{2}\mu + O(\mu^2),\end{aligned} \tag{51}$$

which is consistent with Eqs. (41) and (42).

C. *n* modes with vacuum + *L* probe intensities

Now, we show an estimation of $Y_{11\cdots1}$ in the case that each mode is input with vacuum $+L$ probe intensities. For each mode, the estimation can be given according to Eq. (21). Following a similar method as in the last two sections, we can similarly define the n-mode estimation $Y^{\text{est}}_{11\cdots1}$ of $Y_{11\cdots1}$ according to

$$Y^{\text{est}}_{11\cdots1} = \bigotimes \left(\sum_{j=1}^{\lceil L/2\rceil} \lambda_{2j-1} A_{\mu_{2j-1}} - \sum_{j=1}^{\lfloor L/2\rfloor} \lambda_{2j} A_{\mu_{2j}} + \lambda_0 A_0 \right), \tag{52}$$

where λ_j is defined in Eq. (22). Here, the product $\bigotimes$ denotes a multiplication of A that is defined in Eq. (46). Then the estimation interval is

$$\begin{aligned}\Delta_{L,n} &= \left| \left(\mu_1\mu_2\cdots\mu_L \sum_{j=1}^{L} \frac{\mu_j^{-2}(e^{\mu_j}-1)}{\prod_{1\leqslant n\leqslant L; n\neq j}(\mu_n-\mu_j)} \right)^n - 1 \right| \\ &= \left| \left\{ (-1)^{L+1}\frac{\mu_1\cdots\mu_L}{(L+1)!} + O\left[\frac{\mu_1\cdots\mu_L\sum\mu_l}{(L+2)!}\right] + 1 \right\}^n - 1 \right| \\ &= \frac{n\mu_1\cdots\mu_L}{(L+1)!} + O\left[\frac{n\mu_1\cdots\mu_L\sum\mu_l}{(L+2)!}\right].\end{aligned} \tag{53}$$

Compared to the estimation error of a single photon given in Eq. (20), we can see that an extra factor n is added when simulating n photons. As the estimation error for a single photon decays exponentially to L, the estimation for n photons is still efficient.

In practice, to get $Y^{\text{est}}_{11\cdots1}$, one has to get $A_{\mu_1\mu_2\cdots\mu_n}$. Suppose, for each mode, there are L probe intensities used; then there are L^n different values $A_{\mu_1\mu_2\cdots\mu_n}$ to be measured. For a small number of n, we can see that the estimation is efficient and accurate. However, the total number of probes scales exponentially with n. Therefore, simulating a large number of single photons with a phase-randomized coherent state is not efficient.

V. TOTAL ERROR OF ESTIMATION

In Secs. III D and III E, we showed the total error of the estimation when considering a finite sample size. In general, the total error with n input modes consists of the estimation error $\Delta_{L,n}$ and the statistical error $\Delta_{s,n}$,

$$\Delta_{t,n} \approx \Delta_{s,n} + \Delta_{L,n}. \tag{54}$$

The estimation error $\Delta_{L,n}$ is given in Eq. (53). When Δ_L is small enough and n is not large, we can approximate $\Delta_{L,n}$ by

$$\Delta_{L,n} = n\Delta_{L,1}. \tag{55}$$

The statistical error $\Delta_{s,n}$ consists of a statistical fluctuation when estimating $A_{\mu_1\mu_2\cdots\mu_n}$ for different probe intensities $\{\mu_1\mu_2\cdots\mu_n\}$. Like for the case with one input mode, we consider the same statistical error for all $A_{\mu_1\mu_2\cdots\mu_n}$ using

$$\Delta_{s,n}\left(A_{\mu_1\mu_2\cdots\mu_n}\right) =\lesssim \frac{1}{\sqrt{m}} = \sqrt{\frac{(L+1)^n}{M}}, \tag{56}$$

where M denotes the total number of samples. Note that the estimation $Y^{\text{est}}_{11\cdots1}$ given in Eq. (52) can be reformulated by

$$Y^{\text{est}}_{11\cdots1} = \sum_{j_1,j_2,\ldots,j_n=0}^{L} \lambda_{j_1,j_2,\ldots,j_n} A_{\mu_{j_1}\mu_{j_2}\cdots\mu_{j_n}}, \tag{57}$$

where $\lambda_{j_1,j_2,\ldots,j_n} = \lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_n}$. In this case, the sample error of $Y^{\text{est}}_{11\cdots1}$ can be given by

$$\Delta_{s,n}\left(Y^{\text{est}}_{11\cdots1}\right) = \Delta_{s,n}\left(A_{\mu_{j_1}\mu_{j_2}\cdots\mu_{j_n}}\right) f(L,n), \tag{58}$$

where

$$\begin{aligned} f(L,n) &= \sqrt{\sum_{j_1,j_2,\ldots,j_n=0}^{L} \lambda^2_{j_1,j_2,\ldots,j_n}} \\ &= \sqrt{\sum_{j_1,j_2,\ldots,j_n=0}^{L} \lambda^2_{j_1}\lambda^2_{j_2}\cdots\lambda^2_{j_n}} \\ &= \sqrt{\sum_{j_1=0}^{L}\lambda^2_{j_1}\sum_{j_2=0}^{L}\lambda^2_{j_2}\cdots\sum_{j_n=0}^{L}\lambda^2_{j_n}} \\ &= f(L,1)^n. \end{aligned} \tag{59}$$

Suppose the probe intensities for each mode are $\mu_j = j/L$ for $j = 0,1,\ldots,L$; then we have that

$$\Delta_{t,n} = \sqrt{\frac{(L+1)^n}{M}} f(L,1)^n + n\Delta_{L,1}. \tag{60}$$

Note that we have $\ln\Delta_{L,1} = -2.772L + 3.718$ and $\ln f(L,1) = 0.67L + 0.189$; then

$$\Delta_{t,n} = \sqrt{\frac{(L+1)^n}{M}} e^{n(0.67L+0.189)} + ne^{-2.772L+3.718}. \tag{61}$$

We further optimize over L to get a minimum total error of estimation, as shown in Fig. 7. The optimal number of probe intensities for different input modes is shown in Fig. 8.

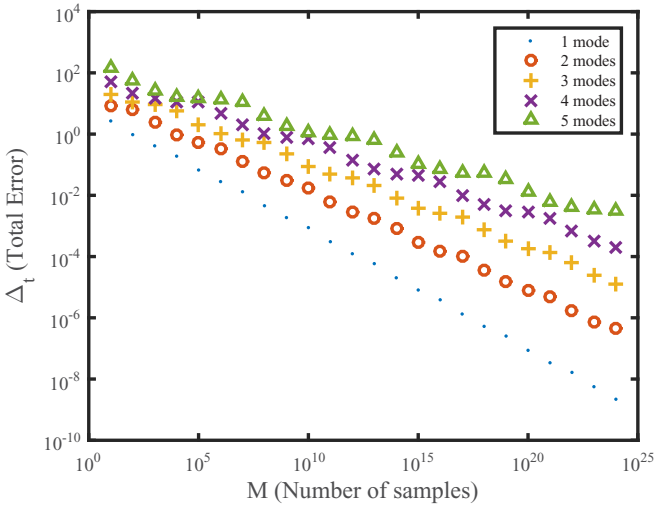

FIG. 7. Optimized total error for different numbers of samples and different input modes.

VI. DISCUSSION

In this work, we propose a way to simulate a single photon with imperfect photon sources. We show that for a single photon, we can efficiently simulate it with a coherent state. In addition, we generalize our result to multiple photon scenarios.

Our result indicates that a small number of single photons can be well simulated by practical photon sources. In practice, this is useful for several information tasks. For instance, in quantum key distribution and quantum random number generation [21,22], we can use phase-randomized coherent states as a source and at the same time guarantee the security. In multipartite measurement-device-independent QKD [23], our results can be applied to increase the key rate. In computation tasks, such as boson sampling, we can simulate the circuit by inputting an imperfect photon source. In the measurement-device-independent entanglement witness for multipartite quantum states [24–26], our method can also make use of an imperfect photon source instead of a single photon to witness multipartite entanglement. In general, our method can also be regarded as a simulator for general quantum computing circuits.

Recently, Valente and Lezama realized our result by experimentally reconstructing the tomography of single-photon temporal states [27]. In both our theoretical scheme and the experimental demonstration, we take the phase-randomized coherent states as an example. It is not hard to see that other practical photon sources with a different mixing of Fock states can also be used in our method. Different photon sources will have similar estimations and errors and may provide some further advantages for certain tasks. Theoretical and experimental investigations of simulating single photons with other practical photon sources are interesting future works.

ACKNOWLEDGMENTS

We acknowledge D. Berry and Z. Cao for insightful discussions. This work was supported by the 1000 Youth Fellowship program in China and the NSERC Discovery Grant.

APPENDIX A: VACUUM + TWO PROBE INTENSITIES

In this case, Alice inputs three phase-randomized coherent states. From the previous calculation, we know that the interval

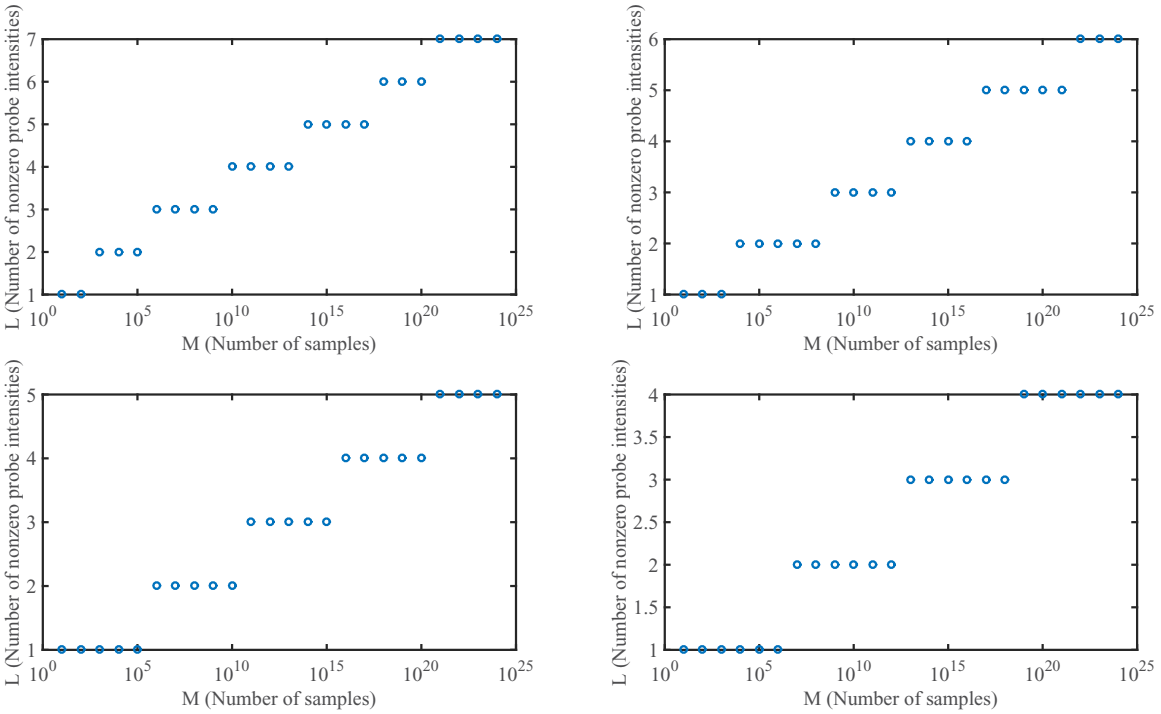

FIG. 8. Optimized number of nonzero probe intensities for different numbers of input modes. Top left, two modes; top right, three modes; bottom left, four modes; and bottom right, five modes.

is minimized when one of the intensities is zero. By assuming that and using two other nonzero intensities, μ,ν, we have three linear relations:

$$\begin{aligned} A_0 &= Y_0, \\ A_\mu &= Y_0 + \mu Y_1 + \frac{\mu^2}{2} Y_2 + \frac{\mu^3}{3!} Y_3 + \cdots, \\ A_\nu &= Y_0 + \nu Y_1 + \frac{\nu^2}{2} Y_2 + \frac{\nu^3}{3!} Y_3 + \cdots. \end{aligned} \tag{A1}$$

First, eliminating Y_0 and defining $B_\mu = (A_\mu - Y_0)/\mu$ and $B_\nu = (A_\nu - Y_0)/\nu$, we get

$$\begin{aligned} \mu B_\mu &= \mu Y_1 + \frac{\mu^2}{2} Y_2 + \frac{\mu^3}{3!} Y_3 + \cdots, \\ \nu B_\nu &= \nu Y_1 + \frac{\nu^2}{2} Y_2 + \frac{\nu^2}{3!} Y_3 + \cdots. \end{aligned} \tag{A2}$$

Then, we can eliminate Y_2,

$$\begin{aligned} \nu^{-1} B_\nu - \mu^{-1} B_\mu = \frac{\mu - \nu}{\mu\nu} Y_1 + \left(\frac{\nu}{3!} Y_3 + \frac{\nu^2}{4!} Y_4 + \cdots \right) \\ - \left(\frac{\mu}{3!} Y_3 + \frac{\mu^2}{4!} Y_4 + \cdots \right), \end{aligned} \tag{A3}$$

and we can estimate Y_1 by

$$Y_1^{\text{est}} = \mu\nu \frac{\nu^{-1} B_\nu - \mu^{-1} B_\mu}{\mu - \nu} \tag{A4}$$

and

$$Y_1 = Y_1^{\text{est}} + \frac{\mu\nu}{\mu - \nu} \left(\frac{\mu - \nu}{3!} Y_3 + \frac{\mu^2 - \nu^2}{4!} Y_4 + \cdots \right). \tag{A5}$$

The estimation interval is given by the difference between the maximal and minimal values of $\frac{\mu\nu}{\mu-\nu}(\frac{\mu-\nu}{3!} Y_3 + \frac{\mu^2-\nu^2}{4!} Y_4 + \cdots)$, that is,

$$\Delta_2 = \frac{\mu\nu}{\mu - \nu} \left(\frac{e^\mu - 1 - \mu}{\mu^2} - \frac{e^\nu - 1 - \nu}{\nu^2} \right). \tag{A6}$$

For small μ and ν, we can approximate Δ_2 by

$$\Delta_2 = \frac{\mu\nu}{3!} + O\left[\frac{\mu\nu(\mu + \nu)}{4!} \right]. \tag{A7}$$

APPENDIX B: DERIVING Y_1^{est} AND Δ_L FOR VACUUM PLUS L PROBE INTENSITIES

Suppose $\mu_0 = 0$ and $\mu_1 < \mu_2 < \cdots < \mu_L$; similar to Eq. (A1), the set of linear equations can be expressed according to

$$\begin{pmatrix} A_0 \\ A_{\mu_1} \\ A_{\mu_2} \\ \vdots \\ A_{\mu_L} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 1 & \mu_1 & \frac{\mu_1^2}{2!} & \frac{\mu_1^3}{3!} & \cdots \\ 1 & \mu_2 & \frac{\mu_2^2}{2!} & \frac{\mu_2^3}{3!} & \cdots \\ & & \vdots & & \\ 1 & \mu_L & \frac{\mu_L^2}{2!} & \frac{\mu_L^3}{3!} & \cdots \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_1 \\ Y_2 \\ \vdots \end{pmatrix}. \tag{B1}$$

We can eliminate the vacuum term by defining $B_{\mu_l} = \frac{A_{\mu_l} - A_0}{\mu_l}$ for $1 \leqslant l \leqslant L$. Then the linear equations becomes

$$\begin{pmatrix} B_{\mu_1} \\ B_{\mu_2} \\ \vdots \\ B_{\mu_L} \end{pmatrix} = \begin{pmatrix} 1 & \mu_1 & \mu_1^2 & \cdots \\ 1 & \mu_2 & \mu_2^2 & \cdots \\ & \vdots & & \\ 1 & \mu_L & \mu_L^2 & \cdots \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2/2! \\ \vdots \end{pmatrix}. \tag{B2}$$

Define $\mathbf{A} = (B_{\mu_1}, B_{\mu_2}, \ldots, B_{\mu_L})^{\mathrm{T}}$, $\mathbf{Y} = (Y_1, Y_2/2!, \ldots, Y_L/L!, \ldots)^{\mathrm{T}}$, and

$$\mathbf{V} = \begin{pmatrix} 1 & \mu_1 & \mu_1^2 & \cdots \\ 1 & \mu_2 & \mu_2^2 & \cdots \\ & \vdots & & \\ 1 & \mu_L & \mu_L^2 & \cdots \end{pmatrix}. \tag{B3}$$

Then, we can rewrite the linear equations as

$$\mathbf{A} = \mathbf{V}\mathbf{Y}, \tag{B4}$$

Define V' to be the first L columns of V,

$$\mathbf{V}' = \begin{pmatrix} 1 & \mu_1 & \mu_1^2 & \cdots \mu_1^{L-1} \\ 1 & \mu_2 & \mu_2^2 & \cdots \mu_2^{L-1} \\ & \vdots & & \\ 1 & \mu_L & \mu_L^2 & \cdots \mu_L^{L-1} \end{pmatrix}. \tag{B5}$$

Then we can see that $\mathbf{V}'$ is a Vandermonde matrix. Denote the inverse of $\mathbf{V}'$ by $\mathbf{M}$, whose element $M_{i,j}$ is given by [28]

$$\begin{aligned} M_{i,j} &= \frac{(-1)^{i-1} \sum_{1 \leqslant k_1 < k_2 < \cdots < k_{L-i} \leqslant L; k_1, k_2, \ldots, k_{L-i} \neq j} \mu_{k_1} \mu_{k_2} \cdots \mu_{k_{L-i}}}{\prod_{1 \leqslant l \leqslant L; l \neq j} (\mu_l - \mu_j)}, \quad 1 \leqslant i < L, \\ M_{n,j} &= \frac{1}{\prod_{1 \leqslant l \leqslant L; l \neq j} (\mu_j - \mu_l)}. \end{aligned} \tag{B6}$$

Then we can multiply $\mathbf{M}$ for both sides of Eq. (B4) and get

$$\mathbf{M}\mathbf{A} = \mathbf{M}\mathbf{V}\mathbf{Y}. \tag{B7}$$

That is,

$$M \begin{pmatrix} B_{\mu_1} \\ B_{\mu_2} \\ \vdots \\ B_{\mu_L} \end{pmatrix} = M \begin{pmatrix} 1 & \mu_1 & \mu_1^2 & \cdots \\ 1 & \mu_2 & \mu_2^2 & \cdots \\ & \vdots & & \\ 1 & \mu_L & \mu_L^2 & \cdots \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2/2! \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots 0 & \sum_{1 \leqslant j \leqslant L} M_{1,j} \mu_j^L & \cdots \\ 0 & 1 & 0 & \cdots 0 & \sum_{1 \leqslant j \leqslant L} M_{2,j} \mu_j^L & \cdots \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots 1 & \sum_{1 \leqslant j \leqslant L} M_{L,j} \mu_j^L & \cdots \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2/2! \\ \vdots \end{pmatrix}. \tag{B8}$$

By considering the first row, we have

$$\sum_{1\leqslant j\leqslant L} M_{1,j}B_{\mu_j} = Y_1 + \sum_{k>L}\frac{Y_k}{k!}\sum_{1\leqslant j\leqslant L} M_{1,j}\mu_j^k, \quad \text{(B9)}$$

where $M_{1,j}$ is given by Eq. (B6),

$$M_{1,j} = \prod_{1\leqslant l\leqslant L; l\neq j}\frac{\mu_l}{\mu_l-\mu_j}. \quad \text{(B10)}$$

Therefore, the estimation of Y_1 is given by

$$\begin{aligned} Y_1^{\text{est}} &= \sum_{1\leqslant j\leqslant L} M_{1,j}B_{\mu_j} \\ &= \sum_{1\leqslant j\leqslant L}\frac{A_{\mu_j}-A_0}{\mu_j}\prod_{1\leqslant l\leqslant L; l\neq j}\frac{\mu_l}{\mu_l-\mu_j} \\ &= \mu_1\mu_2\cdots\mu_L\sum_{j=1}^{L}\frac{\mu_j^{-2}\left(A_{\mu_j}-A_0\right)}{\prod_{1\leqslant l\leqslant L; l\neq j}(\mu_l-\mu_j)}. \end{aligned} \quad \text{(B11)}$$

Defining the remaining term by $R = \sum_{k>L}\frac{Y_k}{k!}\sum_{1\leqslant j\leqslant L} M_{1,j}\mu_j^k$, the estimation interval is given by the interval of the maximal and minimal possible values of R,

$$\Delta_L = \max_{Y_k,\forall k>L} R - \min_{Y_k,\forall k>L} R. \quad \text{(B12)}$$

Denoting $\alpha_j^k = M_{1,j}\mu_j^k$, that is,

$$\alpha_j^k = \mu_j^k\prod_{1\leqslant l\leqslant L; l\neq j}\frac{\mu_l}{\mu_l-\mu_j}, \quad \text{(B13)}$$

we can easily verify that (1) α_j^k is positive when j is odd and (2) $|\alpha_j^k| < |\alpha_{j'}^k|$ when $j < j'$. Therefore, the term $\sum_{1\leqslant j\leqslant k} M_{1,j}\mu_j^k$ can be expressed as

$$\sum_{1\leqslant j\leqslant L} M_{1,j}\mu_j^k = \sum_{1\leqslant j\leqslant L}(-1)^{j-1}\left|\alpha_j^k\right|. \quad \text{(B14)}$$

When L is even, the sum can be grouped into $(|\alpha_1^k| - |\alpha_2^k|) + (|\alpha_3^k| - |\alpha_4^k|) + \cdots + (|\alpha_{L-1}^k| - |\alpha_L^k|)$, and we can see that $\sum_{1\leqslant j\leqslant L} M_{1,j}\mu_j^k$ is negative. When L is odd, the sum can be grouped into $|\alpha_1^k| + (-|\alpha_2^k| + |\alpha_3^k|) + (-|\alpha_4^k| + |\alpha_5^k|) + \cdots + (-|\alpha_{L-1}^k| + |\alpha_L^k|)$, and we can see that $\sum_{1\leqslant j\leqslant L} M_{1,j}\mu_j^k$ is positive. Therefore, the signs of $\sum_{1\leqslant j\leqslant L} M_{1,j}\mu_j^k$ are the same for a fixed L, i.e., $(-1)^{L+1}$. Consequently, the maximum (minimum) value of R can be obtained when all Y_k are equal to the same value (either 0 or 1). We denote those two values as $R_{Y_k=0,\forall k>L}$ and $R_{Y_k=1,\forall k>L}$, respectively.

Defining $R' = (-1)^{L+1}R$, the estimation interval is given by

$$\Delta_L = R'_{Y_k=1,\forall k>L} - R'_{Y_k=0,\forall k>L}. \quad \text{(B15)}$$

Note that $R'_{Y_k=0,\forall k>L} = 0$. To calculate $R'_{Y_k=1,\forall k>L}$, we know that R contains only the $Y_k, \forall k > L$ terms, and therefore, the values of $Y_j, \forall j = 0,1,\ldots,L$ cannot affect the value of R. To simplify the calculation of $A_{\mu_j}, \forall j = 0,1,\ldots,L$, we can consider the case where $Y_0 = Y1 = \cdots = 1$ and hence $A_{\mu_j} = e^{\mu_j}$. In this case, according to Eq. (B9), we have

$$\begin{aligned} \Delta_L &= R'_{Y_k=1,\forall k>L} \\ &= (-1)^{L+1}\sum_{k>L}\frac{1}{k!}\sum_{1\leqslant j\leqslant L} M_{1,j}\mu_j^k \\ &= (-1)^{L+1}\left(Y_1^{\text{est}} - Y_1\right) \\ &= (-1)^{L+1}\left(\mu_1\mu_2\cdots\mu_L\sum_{j=1}^{L}\frac{\mu_j^{-2}(e^{\mu_j}-1)}{\prod_{1\leqslant n\leqslant L; n\neq j}(\mu_n-\mu_j)} - 1\right). \end{aligned} \quad \text{(B16)}$$

The estimation Y_1^{est} in Eq. (B11) can be represented as a linear combination of A_{μ_j} as

$$\begin{aligned} Y_1^{\text{est}} &= \sum_{j=1}^{\lceil L/2\rceil}\frac{\mu_1\mu_2\cdots\mu_L\mu_{2j+1}^{-2}\left(A_{\mu_{2j+1}}-A_0\right)}{\prod_{1\leqslant n\leqslant L; n\neq j}(\mu_n-\mu_{2j+1})} \\ &\quad - \sum_{j=1}^{\lfloor L/2\rfloor}\frac{-\mu_1\mu_2\cdots\mu_L\mu_{2j}^{-2}\left(A_{\mu_{2j}}-A_0\right)}{\prod_{1\leqslant n\leqslant L; n\neq j}(\mu_n-\mu_{2j})} \\ &= \sum_{j=1}^{\lceil L/2\rceil}\lambda_{2j-1}\left(A_{\mu_{2j-1}}-A_0\right) - \sum_{j=1}^{\lfloor L/2\rfloor}\lambda_{2j}\left(A_{\mu_{2j}}-A_0\right) \\ &= \sum_{j=1}^{\lceil L/2\rceil}\lambda_{2j-1}A_{\mu_{2j-1}} - \sum_{j=1}^{\lfloor L/2\rfloor}\lambda_{2j}A_{\mu_{2j}} + \lambda_0 A_0, \end{aligned} \quad \text{(B17)}$$

where the coefficients λ_j are positive and given by

$$\begin{aligned} \lambda_0 &= \sum_{j=1}^{L}(-1)^j\lambda_j, \\ \lambda_j &= \frac{(-1)^{j+1}}{\mu_j}\prod_{1\leqslant n\leqslant L; n\neq j}\frac{\mu_n}{(\mu_n-\mu_j)}, \quad 1\leqslant j\leqslant L. \end{aligned} \quad \text{(B18)}$$

Similarly, the estimation error Δ_L is given by

$$\Delta_L = (-1)^{L+1}\left(\sum_{j=1}^{\lceil L/2\rceil}\lambda_{2j-1}e^{\mu_{2j-1}} - \sum_{j=1}^{\lfloor L/2\rfloor}\lambda_{2j}e^{\mu_{2j}} + \lambda_0 - 1\right). \quad \text{(B19)}$$

[1] P. W. Shor, SIAM J. Comput. **26**, 1484 (1997).

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

[3] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

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

[5] J. Yin, J.-G. Ren, H. Lu, Y. Cao, H.-L. Yong, Y.-P. Wu, C. Liu, S.-K. Liao, F. Zhou, Y. Jiang, X.-D. Cai, P. Xu, G.-S. Pan, J.-J. Jia, Y.-M. Huang, H. Yin, J.-Y. Wang, Y.-A. Chen, C.-Z. Peng, and J.-W. Pan, Nature (London) **488**, 185 (2012).
[6] N. J. Cerf, C. Adami, and P. G. Kwiat, Phys. Rev. A **57**, R1477(R) (1998).
[7] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, Rev. Mod. Phys. **79**, 135 (2007).
[8] E. Knill, R. Laflamme, and G. J. Milburn, Nature (London) **409**, 46 (2001).
[9] D. Gottesman and I. L. Chuang, Nature (London) **402**, 390 (1999).
[10] X.-C. Yao, T.-X. Wang, P. Xu, H. Lu, G.-S. Pan, X.-H. Bao, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, and J.-W. Pan, Nat. Photonics **6**, 225 (2012).
[11] Y.-F. Huang, B.-H. Liu, L. Peng, Y.-H. Li, L. Li, C.-F. Li, and G.-C. Guo, Nat. Commun. **2**, 546 (2011).
[12] G. Brassard, N. Lütkenhaus, T. Mor, and B. C. Sanders, Phys. Rev. Lett. **85**, 1330 (2000).
[13] W.-Y. Hwang, Phys. Rev. Lett. **91**, 057901 (2003).
[14] H.-K. Lo, X. Ma, and K. Chen, Phys. Rev. Lett. **94**, 230504 (2005).
[15] X.-B. Wang, Phys. Rev. Lett. **94**, 230503 (2005).
[16] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, Phys. Rev. A **72**, 012326 (2005).
[17] J. W. Harrington, J. M. Ettinger, R. J. Hughes, and J. E. Nordholt, arXiv:quant-ph/0503002.
[18] M. Curty, F. Xu, W. Cui, C. C. W. Lim, K. Tamaki, and H.-K. Lo, Nat. Commun. **5**, 3732 (2014).
[19] F. Xu, H. Xu, and H.-K. Lo, Phys. Rev. A **89**, 052333 (2014).
[20] Z. Zhang, Q. Zhao, M. Razavi, and X. Ma, arXiv:1611.02524 (2016).
[21] Z. Cao, H. Zhou, X. Yuan, and X. Ma, Phys. Rev. X **6**, 011020 (2016).
[22] X. Yuan, H. Zhou, Z. Cao, and X. Ma, Phys. Rev. A **92**, 022124 (2015).
[23] Y. Fu, H.-L. Yin, T.-Y. Chen, and Z.-B. Chen, Phys. Rev. Lett. **114**, 090501 (2015).
[24] C. Branciard, D. Rosset, Y.-C. Liang, and N. Gisin, Phys. Rev. Lett. **110**, 060405 (2013).
[25] P. Xu, X. Yuan, L.-K. Chen, H. Lu, X.-C. Yao, X. Ma, Y.-A. Chen, and J.-W. Pan, Phys. Rev. Lett. **112**, 140506 (2014).
[26] Q. Zhao, X. Yuan, and X. Ma, Phys. Rev. A **94**, 012343 (2016).
[27] P. Valente and A. Lezama, arXiv:1610.06854.
[28] N. Macon and A. Spitzbart, Am. Math. Mon. **65**, 95 (1958).