Differential-phase-shift quantum-key-distribution protocol with a small number of random delays

Yuki Hatakeyama,[1] Akihiro Mizutani,[1,*] Go Kato,[2] Nobuyuki Imoto,[1] and Kiyoshi Tamaki[3]
[1]*Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan*
[2]*NTT Communication Science Laboratories, NTT Corporation, 3-1 Morinosato Wakamiya, Atsugi, Kanagawa 243-0198, Japan*
[3]*NTT Basic Research Laboratories, NTT Corporation, 3-1 Morinosato Wakamiya, Atsugi, Kanagawa 243-0198, Japan*
(Received 18 January 2017; published 3 April 2017)

The differential-phase-shift (DPS) quantum-key-distribution (QKD) protocol was proposed aiming at simple implementation, but it can tolerate only a small disturbance in a quantum channel. The round-robin DPS (RRDPS) protocol could be a good solution for this problem, which in fact can tolerate even up to 50% of a bit error rate. Unfortunately, however, such a high tolerance can be achieved only when we compromise the simplicity, i.e., Bob's measurement must involve a large number of random delays ($|\mathcal{R}|$ denotes its number), and in a practical regime of $|\mathcal{R}|$ being small, the tolerance is low. In this paper, we propose a DPS protocol to achieve a higher tolerance than the one in the original DPS protocol, in which the measurement setup is less demanding than the one of the RRDPS protocol for the high tolerance regime. We call our protocol the small-number-random DPS (SNRDPS) protocol, and in this protocol, we add only a small amount of randomness to the original DPS protocol, i.e., $2 \leqslant |\mathcal{R}| \leqslant 10$. In fact, we found that the performance of the SNRDPS protocol is significantly enhanced over the original DPS protocol only by employing a few additional delays such as $|\mathcal{R}| = 2$. Also, we found that the key generation rate of the SNRDPS protocol outperforms the RRDPS protocol without monitoring the bit error rate when it is less than 5% and $|\mathcal{R}| \leqslant 10$. Our protocol is an intermediate protocol between the original DPS protocol and the RRDPS protocol, and it increases the variety of the DPS-type protocols with quantified security.

DOI: 10.1103/PhysRevA.95.042301

I. INTRODUCTION

Quantum key distribution (QKD) holds promise for realizing information-theoretically secure communication between two distant parties (Alice and Bob) against any eavesdropper (Eve). Since the first invention of the BB84 protocol [1], many QKD protocols have been proposed [2–7]. Among them, the differential-phase-shift (DPS) QKD [5] can be rather simply implemented with a passive detection unit. A field demonstration of the DPS protocol [8] has been already conducted, and the information-theoretical security proof of the DPS protocol has been established by Tamaki *et al.* [9,10]. Unfortunately, however, this proof shows that the DPS protocol can tolerate only a small bit error rate regime (less than 4% with a typical block length of L light pulses, say $L = 32$).

Recently, in order to solve this problem, a new type of protocol called the round-robin differential-phase-shift (RRDPS) QKD protocol [11] was proposed. This is a modified protocol from the original DPS protocol in that Bob's measurement has a freedom to randomly choose which pair of the incoming pulses to be interfered. This modification brings a distinct feature to the RRDPS protocol that the security can be guaranteed without monitoring any disturbance between Alice and Bob. Moreover, when the number of random delays (we denote it by $|\mathcal{R}|$) is large, the RRDPS protocol has a strong tolerance to the bit error rate, and surprisingly it can tolerate the bit error rate of even 50% when $|\mathcal{R}| \to \infty$. Thanks to these features, the RRDPS protocol has attracted theoretical works [12–16], and proof-of-principle experiments have been demonstrated [12,17–19]. Unfortunately, however, an experimental implementation of the RRDPS protocol is not as simple as the one of the original DPS protocol. One of the main technological challenges for its realization is to switch the delay at random for each block of the pulses at Bob's measurement. According to the number of the random delays, some passive interferometers [12] or a variable-delay interferometer [17] or some optical switches [18] are needed for Bob's measurement. Obviously, when the number of the random delays increases, an implementation of Bob's measurement will be complicated. Hence, from a practical viewpoint, it is preferable to implement the RRDPS protocol with small $|\mathcal{R}|$, for instance, $|\mathcal{R}| = 4$ as demonstrated in [12].

In this paper, we consider improving the bit error tolerance of the DPS protocol without significantly increasing the difficulties of its implementation. For this, we consider adding only a small amount of randomness, say $2 \leqslant |\mathcal{R}| \leqslant 10$, to the DPS protocol. This modification can also be seen as the modification from the RRDPS protocol in that our protocol exploits more pulses than that of the RRDPS protocol for a given $|\mathcal{R}|$. Importantly, this modification does not increase any experimental difficulty at Bob's side. We call our protocol the small-number-random DPS (SNRDPS) protocol. We present the information-theoretical security of our protocol, in which we have made some assumptions on the devices. In particular, we assume perfect phase modulations (namely, Alice's phase modulation is exactly 0 or π) and block-wise phase randomization (the state of the L pulses is a classical mixture of photon number states). With these assumptions, we prove the security based on the Shor-Preskill's security proof [20]. By using the result of the security proof, we compare the performance of the SNRDPS protocol and the one of the original DPS protocol. As a result, we found that the key generation rate is significantly improved only with a few additional delays, say $|\mathcal{R}| = 2$. For instance, if the bit error rate $e^{(b)}$ is 2%, the key generation rate of the SNRDPS protocol with $|\mathcal{R}| = 2$ scales as $O(\eta^{3/2})$

*mizutani@qi.mp.es.osaka-u.ac.jp

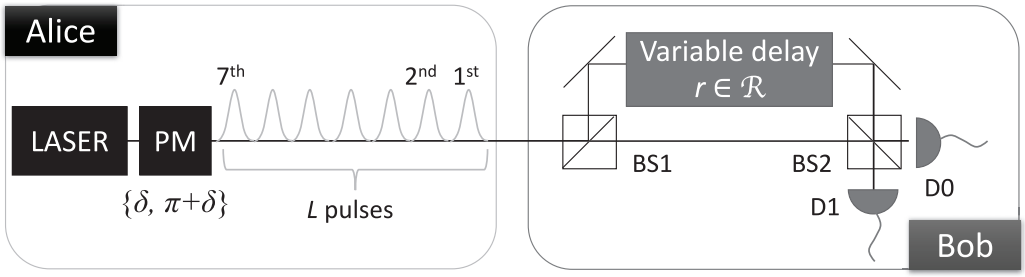


FIG. 1. Schematics of the SNRDPS protocol for $L = 7$ and $\mathcal{R} = \{1,6,2,5\}$. Alice sends seven coherent pulses after she applies a random phase shift δ or $\delta + \pi$ to each of the pulses with a phase modulator (PM), where δ is chosen uniformly and at random from $[0,2\pi)$. After Bob receives the incoming seven pulses, he splits them into two blocks of seven pulses with the first beam splitter (BS1), and shifts backward by rT to one of the blocks. Here, T denotes the interval between two adjacent pulses of the incoming block. By using the second beam splitter (BS2) and two photon-number-resolving detectors (D0 and D1), he observes the relative phase of two pulses in the block. Note that, each bit in the sifted key is generated from a block where Bob has detected one photon from one pair of the interfering pulses and has detected the vacuum in all the other pulses. We call such an event *detected* event. For example, if $r = 2$ is chosen and Bob detects exactly one photon in the pair of (3^{rd}, 5^{th}) interfering pulses and detects the vacuum in all of the other pulses, he obtains the relative phase between the 3^{rd} and the 5^{th} interfering pulses. If the relative phase is $0(\pi)$, he obtains the sifted bit $s_{\text{B}} = 0(1)$.

with a channel transmittance η in a longer distance regime while the original DPS protocol scales as $O(\eta^2)$. Also, when $e^{(\text{b})} = 5\%$, the SNRDPS protocol with $|\mathcal{R}| = 2$ gives a positive key generation rate while the original DPS protocol cannot give a positive one. Therefore, the small number of random delays provides a significant improvement in the resulting key generation rate compared to the one of the DPS protocol.

Moreover, we compare the key generation rates of the SNRDPS protocol and the ones of the RRDPS protocol without monitoring the disturbance when the same number of the random delays $|\mathcal{R}|$ is employed between two protocols. Consequently, we found that the SNRDPS protocol gives a better key generation rate than the one of the RRDPS protocol without monitoring the bit error rate when $|\mathcal{R}|$ is less than 10 and the bit error rate is small such as less than 5%.

This paper is organized as follows. First, in Sec. II we explain the DPS-type QKD protocol including the assumptions on Alice and Bob's devices. Next, in Sec. III we prove the security of the DPS-type protocol, where our security proof is based on the Shor-Preskill's security proof [20]. After that, in Sec. IV we show the simulation results for the SNRDPS protocol, and compare the key generation rates with the various numbers of the random delays $|\mathcal{R}| = \{2,4,6,8,10\}$. Finally, we summarize the paper in Sec. V.

II. DPS-TYPE QKD PROTOCOL

In this section, before providing the description of the actual protocol, we first list up the assumptions on Alice and Bob's devices. See Fig. 1 for the actual setup.

A. Assumptions on Alice and Bob's devices

First, we describe the assumptions on Alice's source. We assume that it emits a single-mode coherent light pulse, and Alice splits its pulse into a block of L pulses. The L pulses are block-wise phase randomized, namely, the quantum state of the L pulses is described as a classical mixture of photon number states. The relative phase between the adjacent pulses is modulated by 0 or π according to her randomly chosen bit 0 or 1, respectively.

Next, as for Bob's device, it is equipped with two photon-number-resolving (PNR) detectors that can discriminate among 0, 1, and more than 1 photon. He first splits L incoming pulses into two blocks of L pulses by using a 50:50 beam splitter (BS), shifts backward only one of the L-pulse blocks by r that is chosen randomly from the set $\mathcal{R} \subset \{1,2,\ldots,L-1\}$. Then, the first $L-r$ pulses in the shifted block will be interfered with the last $L-r$ pulses in the other block with another $50:50$ BS, and then Bob performs a photon measurement with the PNR detectors. Each of the detectors corresponds to the bit value of 0 and 1, respectively (see Fig. 1). Finally, we assume that there is no side channel.

B. DPS-type QKD

We describe the "DPS-type" QKD protocol, which is the generalization of the DPS QKD protocol in that it employs the arbitrary number of random delays, and therefore the DPS-type protocol includes both the original DPS protocol and the RRDPS protocol. The protocol of the DPS-type QKD runs as follows.

(A1) Alice generates a random L-bit string $\vec{s} \equiv (s_1,s_2,\ldots,s_L)$, a random number $\delta \in [0,2\pi)$, and then she prepares a block of L coherent pulses in the following state:

$$|\Psi\rangle = \bigotimes_{k=1}^{L} |(-1)^{s_k} e^{i\delta}\alpha\rangle_k , \tag{1}$$

where $|(-1)^{s_k} e^{i\delta}\alpha\rangle_k$ represents the coherent state of the k^{th} pulse. She sends $|\Psi\rangle$ to Bob through a quantum channel.

(A2) Bob splits the incoming L pulses into two L-pulse blocks by using the 50:50 BS. He applies a delay rT to one of the paths in the Mach-Zehnder interferometer, where T denotes the interval between two adjacent pulses in the block and r is chosen uniformly at random from the set $\mathcal{R} \subset \{1,2,\ldots,L-1\}$. After that, Bob makes interference between two L-pulse blocks by using the other 50:50 BS and performs the photon detection with the photon number resolving detectors. Let us call the event *detected* if he detects exactly one photon in the pair of (k^{th}, $(k+r)^{\text{th}}$) (with

$1 \leqslant k \leqslant L-r$) interfering pulses and detects the vacuum in all of the other pulses (including $2r$ half pulses that do not interfere with any other half pulses). If the event is not *detected*, Alice and Bob skip steps (A3) and (A4).

(A3) Bob takes note of the detected bit value s_{B} and announces the pair of numbers $(i,j)=(k,k+r)$ over an authenticated public channel.

(A4) Alice takes note of the bit value $s_{\mathrm{A}}=s_k\oplus s_{k+r}$.

(A5) Alice and Bob repeat steps (A1)–(A4) N times, and let NQ be the number of the *detected* events.

(A6) Alice and Bob randomly select a small portion ξ of NQ *detected* events, and compare the bit values over an authenticated public channel. This gives the estimate of the bit error rate.

(A7) Alice and Bob discuss over an authenticated public channel to perform error correction and privacy amplification on the remaining portion to share a final key of length $GN(1-\xi)$.

Note that the DPS-type protocol includes the original DPS and the RRDPS protocols by choosing $\mathcal{R}$ in step (A2) as $\mathcal{R}=\{1\}$ and $\mathcal{R}=\{1,2,\ldots,L-1\}$, respectively. Also, we define the SNRDPS protocol by setting $\mathcal{R}=\bigcup_{m=1}^{t}\{m,L-m\}$ with $0<t<L/2$.

III. SECURITY PROOF

In this section, we prove the security of the SNRDPS protocol with $\mathcal{R}=\bigcup_{m=1}^{t}\{m,L-m\}$ for $0<t<L/2$. Our security proof can be summarized as follows. First, in Sec. III A we convert the actual protocol to an *alternative* protocol for simplicity of the analysis, where Bob performs the alternative measurement (we call it the dial measurement) instead of the actual one. Note that, by switching Bob's actual measurement with the delays r and $L-r$ uniformly at random, we show that he can simulate the dial measurement characterized by the delay r with a 50% of additional detection loss (see Lemma 1 below). Therefore, by introducing the additional loss in the alternative measurement, the dial measurement is equivalent to the actual measurement, and therefore we can employ the alternative measurement in the security proof. Next, in Sec. III B we introduce an entanglement distillation protocol as a *virtual* protocol to prove the security of the protocol with the alternative measurement. After that, we construct the POVM elements corresponding to the bit and phase error rates in Sec. III C, and derive a relation between the bit and phase errors by employing some constraint on Alice's sending state in the virtual protocol in Sec. III D, and obtain an upper bound on the phase error rate as the function of the bit error rate in Sec. III E.

A. Bob's alternative measurement

In this subsection, we introduce Bob's alternative measurement, which we will employ in the security proof. In step (A2), Bob extracts the *detected* events, in which only one photon is contained in the incoming L pulses. Here, $\{|k\rangle_{\mathrm{B}}\}_{k=1}^{L}$ denotes the set of basis vectors of the Hilbert space $\mathcal{H}_{\mathrm{B}}$, and $|k\rangle_{\mathrm{B}}$ represents that the k^{th} pulse sent is in a single-photon state. Let $\{\hat{B}_{k,s}^{(r)}\}_{k,s}$ be the POVM for the bit value s detected at the pair of $(k^{\mathrm{th}},(k+r)^{\mathrm{th}})$ interfering pulses under the condition that the delay r is chosen. Considering the effect of the 50:50 BS, $\{\hat{B}_{k,s}^{(r)}\}$ is written as

$$\hat{B}_{k,s}^{(r)}:=\frac{1}{2}\hat{P}\left(\frac{|k\rangle_{\mathrm{B}}+(-1)^{s}|k+r\rangle_{\mathrm{B}}}{\sqrt{2}}\right), \tag{2}$$

for $1\leqslant k\leqslant L-r$. Here, we define $\hat{P}(|\phi\rangle)\equiv|\phi\rangle\langle\phi|$. From Eq. (2), the probability of obtaining the bit value s and the pair of interfering pulses $(k,k+r)$ in his measurement with the delay r is given by $\mathrm{Tr}(\hat{\rho}\hat{B}_{k,s}^{(r)})$ for an arbitrary state $\hat{\rho}$ given that exactly one photon is contained in the L pulses.

Next, for simplicity of the security analysis, we convert Bob's actual measurement into the alternative one. We call it the dial measurement, which gives the relative phase of an arbitrary pair of $(i^{\mathrm{th}},j^{\mathrm{th}})$ (with $i<j$) interfering pulses such that $j-i=r$ or $L-r$ for given r [see Fig. 2(a)]. This measurement has more symmetry than the actual measurement, which makes our analysis much simpler, and importantly it is equivalent to the actual measurement except for 50% of losses as we explain in Lemma 1 below. The POVM $\{\hat{E}_{k,s}^{(r)}\}_{k,s}$ of the dial measurement with the delay r is defined by

$$\hat{E}_{k,s}^{(r)}:=\frac{1}{2}\hat{P}\left(\frac{|k\rangle_{\mathrm{B}}+(-1)^{s}|k+_{L}r\rangle_{\mathrm{B}}}{\sqrt{2}}\right), \tag{3}$$

for $1\leqslant k\leqslant L$. Here, $+_L$ denotes the summation in modulo L, namely, for integers (p,q) with $1\leqslant p\leqslant L$ and $1\leqslant q\leqslant L$,

$$p+_{L}q=\begin{cases}p+q & \text{if } p+q\leqslant L,\\ p+q-L & \text{if } p+q\geqslant L+1.\end{cases} \tag{4}$$

If Bob performs the dial measurement with the delay r, the probability of obtaining the bit value s and the pair of interfering pulses $(k,k+_{L}r)$ is given by $\mathrm{Tr}(\hat{\rho}\hat{E}_{k,s}^{(r)})$. Note that the following relation holds for $\hat{E}_{k,s}^{(r)}$:

$$\hat{E}_{k+_{L}r,s}^{(L-r)}=\hat{E}_{k,s}^{(r)}. \tag{5}$$

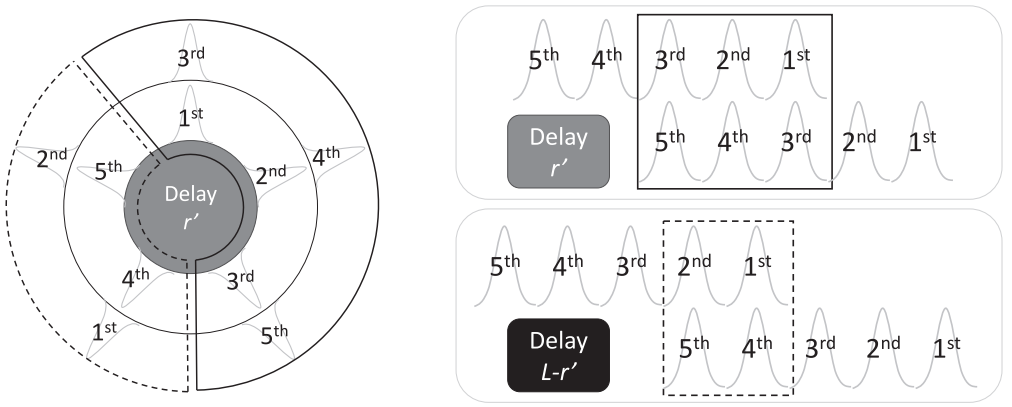
Next, we introduce the following lemma [11] that relates the dial and actual measurements. See Appendix A for its proof.

Lemma 1. We define two conditional probabilities $\Pr[s\wedge(i,j)|r']_{\mathrm{dial}}$ and $\Pr[s\wedge(i,j)|r\in\{r',L-r'\}]_{\mathrm{actual}}$. $\Pr[s\wedge(i,j)|r']_{\mathrm{dial}}$ represents the probability that Bob obtains the bit value s from $(i^{\mathrm{th}},j^{\mathrm{th}})$ (with $i<j$) interfering pulses given that he performs the dial measurement with the delay $r=r'\in\mathcal{R}$. $\Pr[s\wedge(i,j)|r\in\{r',L-r'\}]_{\mathrm{actual}}$ represents the probability that Bob obtains s from $(i^{\mathrm{th}},j^{\mathrm{th}})$ interfering pulses given that he performs the actual measurement with the delays $r=r'$ or $r=L-r'$ chosen uniformly at random. Then, for an arbitrary fixed $r'\in\mathcal{R}$ and for any input state $\hat{\rho}$,

$$\Pr[s\wedge(i,j)|r']_{\mathrm{dial}}=2\Pr[s\wedge(i,j)|r\in\{r',L-r'\}]_{\mathrm{actual}} \tag{6}$$

holds.

Lemma 1 means that the dial measurement with the delay r' after performing a half transmittance filter has the same probability distribution of s and (i,j) as the one of the actual measurement when Bob switches delays r' and $L-r'$ uniformly at random [see Fig. 2(b)]. In other words, Eve cannot distinguish which of the measurements was actually employed from the classical information announced by Bob. Thanks to Eq. (6), we are allowed to use the dial measurement for



(a) Bob's alternative (dial) measurement (b) Bob's actual measurement

FIG. 2. Schematics of (a) Bob's "dial measurement" with $L = 5$ and the delay $r' = 2$, and (b) the corresponding actual measurement. In the dial measurement, there are five patterns of successful detection events. Concretely, he obtains the relative phase between one of the following five pairs of interfering pulses: (1^{st}, 3^{rd}), (2^{nd}, 4^{th}), (3^{rd}, 5^{th}) represented by the regime surrounded by the solid line (1^{st}, 4^{th}), or (2^{nd}, 5^{th}) represented by the regime surrounded by the dashed line. In the actual measurement, by using a specific delay r' or $L - r'$, he can obtain a part of the information that he could obtain in his dial measurement with the delay r'. For example, if he employs the delay $r'(=2)$ in (b), he can obtain the relative phase between either (1^{st}, 3^{rd}), (2^{nd}, 4^{th}) or (3^{rd}, 5^{th}) interfering pulses in (a). These events correspond to those in the regime surrounded by the solid line in (b). Also, if he employs the delay $L - r'(=3)$ in (b), he can obtain the relative phase between either (1^{st}, 4^{th}) or (2^{nd}, 5^{th}) interfering pulses in (a). These events correspond to those in the regime surrounded by the dashed line in (b). Therefore, if he switches the delay r' or $L - r'$ uniformly at random in his actual measurement, he can simulate the dial measurement with the delay r', while the detection efficiency of his actual measurement is half that of his dial measurement (see Lemma 1 below).

proving the security of the actual protocol. We call the protocol where Bob performs the dial measurement instead of the actual one the *alternative* protocol. The alternative protocol runs the same as the actual protocol except for steps (A2) and (A3), which are replaced with the following steps (A2') and (A3'), respectively.

(A2) Bob receives the incoming L pulses and splits them into two L-pulse blocks by using the 50:50 BS. He selects a delay r uniformly at random from the set $\mathcal{R} \subset \{1,2,\ldots,L-1\}$. After that, Bob performs the dial measurement. Let us call the event *detected* if he detects exactly one photon in the pair of interfering pulses $(k^{th},(k +_L r)^{th})$ (with $1 \leqslant k \leqslant L$), and detects the vacuum in all the other pairs of interfering pulses. If the event is not *detected*, Alice and Bob skip steps (A3') and (A4).

(A3) Bob takes note of the detected bit value and announces the pair of numbers $(i,j) = (\min\{k,k +_L r\}, \max\{k,k +_L r\})$ over an authenticated public channel.

B. Virtual protocol

In this subsection, we introduce the entanglement distillation protocol to prove the security of the alternative protocol. Our analysis is based on the Shor-Preskill's security proof [20], where we follow similar arguments of the security proof of the original DPS protocol [9]. To show Alice and Bob virtually extract a maximally entangled state, we need to introduce ancilla systems on Alice's side and decompose Bob's measurement, which we explain below.

First, we explain Alice's sending state in the virtual protocol. Suppose Alice has a quantum register of the L-qubit system and let $\mathcal{H}_A = \bigotimes_{k=1}^{L} \mathcal{H}_{A,k}$ be the Hilbert space of these systems. Then, Alice's state preparation is equivalent to the preparation of the following state over the quantum register system and L pulses as

$$|\Phi_\delta\rangle := 2^{-L/2} \sum_{\vec{s}} \bigotimes_{k=1}^{L} (\hat{H}\,|s_k\rangle_{A,k})\,|(-1)^{s_k} e^{i\delta}\alpha\rangle_k\,, \qquad (7)$$

where $\hat{H} \equiv \frac{1}{\sqrt{2}} \sum_{s,s'=0,1} (-1)^{ss'} |s\rangle\langle s'|$ denotes the Hadamard operator. Note that $\delta \in [0,2\pi)$ is chosen uniformly at random for each preparation of the state $|\Phi_\delta\rangle$. As shown in step (A4), the information that Alice needs to obtain is $s_i \oplus s_j$. To obtain this information, she applies the following quantum circuit (see Fig. 3) to the qubits i and j upon receiving from Bob, and measures the qubit Aq in the computational basis $\{|0\rangle_{Aq}, |1\rangle_{Aq}\}$. The set of measurement operators that Alice

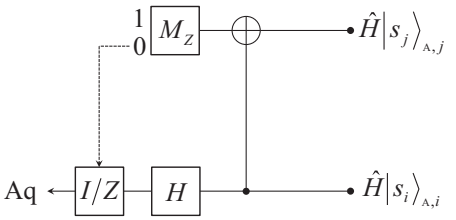


FIG. 3. The quantum circuit representing Alice's procedure in the virtual protocol. The inputs are the i^{th} and j^{th} qubits, where the pair of integers (i,j) (with $i < j$) is announced by Bob. She applies a C-NOT gate (defined on the Z basis $\{|0\rangle_A, |1\rangle_A\}$) to these qubits, and the i^{th} qubit is subjected to a Hadamard gate (denoted by H) while the j^{th} qubit is measured on the Z basis (denoted by M_Z). If the outcome of the Z measurement is 0, she applies a phase flip gate to the i^{th} qubit with probability $1/2$, which is denoted by I/Z. Otherwise, she applies the identity operation to the i^{th} qubit. After that, we name the quantum system of the i^{th} qubit Aq.

performs can be represented by

$$\hat{M}_1^{(i,j)} = \hat{H}\,|0\rangle_{\mathrm{Aq}}\,(_{\mathrm{A},i}\,\langle 0|_{\,\mathrm{A},j}\,\langle 1| + \hat{H}\,|1\rangle_{\mathrm{Aq}}\,(_{\mathrm{A},i}\,\langle 1|_{\,\mathrm{A},j}\,\langle 0|),$$
$$\hat{M}_2^{(i,j)} = \frac{1}{\sqrt{2}}\,|0\rangle_{\mathrm{Aq}}\,(_{\mathrm{A},i}\,\langle 0|_{\,\mathrm{A},j}\,\langle 0| + {}_{\mathrm{A},i}\,\langle 1|_{\,\mathrm{A},j}\,\langle 1|),$$
$$\hat{M}_3^{(i,j)} = \frac{1}{\sqrt{2}}\,|1\rangle_{\mathrm{Aq}}\,(_{\mathrm{A},i}\,\langle 0|_{\,\mathrm{A},j}\,\langle 0| - {}_{\mathrm{A},i}\,\langle 1|_{\,\mathrm{A},j}\,\langle 1|). \tag{8}$$

Note that Alice's state preparation of $|\Phi_\delta\rangle$ with a random and uniform δ followed by the measurement $\{M_m^{(i,j)}\}_{m=1,2,3}$ is equivalent to the step (A1). Moreover, in the virtual protocol, instead of $|\Phi_\delta\rangle$, Alice prepares the following state for simplicity of analysis.

$$|\Phi\rangle := 2^{-L/2}\sum_{\vec{s}}\sum_{\nu=0}^{\infty}|\nu\rangle_{\mathrm{C}}\,\hat{\pi}_\nu\bigotimes_{k=1}^{L}(\hat{H}\,|s_k\rangle_{\mathrm{A},k})\,|(-1)^{s_k}\alpha\rangle_k\,. \tag{9}$$

Here, C is the system that stores the number of photons contained in the L pulses whose Hilbert space is spanned by an orthogonal basis $\{|\nu\rangle_{\mathrm{C}}\}_{\nu=0}^{\infty}$. Also, $\hat{\pi}_\nu$ is the projection onto the subspace that the total photon number in the L pulses is ν. From Eve's perspective, accessible quantum information of Eqs. (7) and (9) are the same since the following equation holds.

$$\frac{1}{2\pi}\int_0^{2\pi} d\delta|\Phi_\delta\rangle\langle\Phi_\delta| = \mathrm{Tr}_{\mathrm{C}}|\Phi\rangle\langle\Phi|. \tag{10}$$

Next, we explain Bob's measurement procedure in the virtual protocol. In principle, he is able to determine whether the event is *detected* or not before he determines the pair of interfering pulses and the bit value by performing the quantum nondemolition (QND) measurement of the total photon number in the incoming L pulses. The event is called *detected* if and only if the measurement outcome of the QND measurement is exactly one photon in the block of L pulses. In the *detected* events, the dial measurement is decomposed into two measurements, namely, the POVM in Eq. (3) is decomposed into

$$\hat{E}_{k,s}^{(r)} = \hat{F}_k^{(r)\dagger}\hat{P}(|s\rangle_{\mathrm{Bq}})\hat{F}_k^{(r)}. \tag{11}$$

Here, the set of measurement operators $\hat{F}_k^{(r)} : \mathcal{H}_{\mathrm{B}} \to \mathcal{H}_{\mathrm{Bq}}$ represents a filtering operation that gives the outcome k and leaves a qubit system $\mathcal{H}_{\mathrm{Bq}}$, which is defined by

$$\hat{F}_k^{(r)} := \begin{cases} \frac{1}{\sqrt{2}}\hat{H}\,|1\rangle_{\mathrm{Bq\,B}}\,\langle k| + \frac{1}{\sqrt{2}}\hat{H}\,|0\rangle_{\mathrm{Bq\,B}}\,\langle k+r| \\ \qquad\qquad \text{if } 1 \leqslant k \leqslant L-r, \\ \frac{1}{\sqrt{2}}\hat{H}\,|1\rangle_{\mathrm{Bq\,B}}\,\langle k+r-L| + \frac{1}{\sqrt{2}}\hat{H}\,|0\rangle_{\mathrm{Bq\,B}}\,\langle k| \\ \qquad\qquad \text{if } L-r+1 \leqslant k \leqslant L. \end{cases} \tag{12}$$

By using Alice's quantum circuit represented by Fig. 3 and Bob's filtering operation described by $\{\hat{F}_k^{(r)}\}_k$, we introduce the following entanglement distillation protocol (EDP).

(V1) Alice prepares $|\Phi\rangle$ and sends a block of L pulses to Bob through a quantum channel.

(V2) Bob receives the incoming L pulses and performs the QND measurement of the total photon number in the L pulses. Let us call the event *detected* if he detects exactly one photon in the block of L pulses. If the event is not *detected*, Alice and Bob skip steps (V3) and (V4) below.

(V3) Bob chooses r uniformly at random from the set $\mathcal{R} = \bigcup_{m=1}^{t}\{m, L-m\}$, and performs the filtering operation $\{\hat{F}_k^{(r)}\}$. He obtains the pair of pulses $(k, k+_L r)$ where exactly one photon is contained, and obtains the output qubit Bq. He sends the pair of integers $(i,j) := (\min\{k, k+_L r\}, \max\{k, k+_L r\})$ to Alice over an authenticated public channel.

(V4) Alice applies the quantum circuit in Fig. 3 on her i^{th} and j^{th} qubits, and outputs the qubit Aq. Also, she measures system C and learns the total photon number ν in the block of L pulses.

(V5) Alice and Bob repeat steps (V1)–(V4) N times. Let NQ be the number of the *detected* events. At this point, Alice and Bob share NQ pairs of qubits.

(V6) Alice and Bob randomly select a small portion ξ of the NQ *detected* events, measure the qubits in the computational basis $\{|0\rangle, |1\rangle\}$, and compare the bit values over the public channel. This gives the estimate of the bit error rate $e^{(\mathrm{b})}$ and hence the number $e^{(\mathrm{b})}NQ(1-\xi)$ of bit errors in the remaining portion.

(V7) Alice and Bob discuss over the public channel to perform entanglement distillation on the remaining pairs of qubits. Finally, they measure all the remaining pairs of qubits on the computational basis to obtain a final key of length $GN(1-\xi)$.

In this protocol, the key generation rate per sending pulse is written as [21]

$$G = [Q(1 - h(e^{(\mathrm{b})})) - h^{(\mathrm{ph})}]/L, \tag{13}$$

where $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$, and $h^{(\mathrm{ph})}$ expresses the number of privacy amplification. The explicit formula of $h^{(\mathrm{ph})}$ is described by

$$h^{(\mathrm{ph})} = \sum_{\nu=0}^{\infty} Q^{(\nu)} h(e^{(\mathrm{ph},\nu)}), \tag{14}$$

where $Q^{(\nu)}$ denotes the function of *detected* events when Alice emits ν photons satisfying $\sum_{\nu=0}^{\infty} Q^{(\nu)} = Q$, and $e^{(\mathrm{ph},\nu)}$ denotes the phase error rate when Alice emits ν photons. Note that the equivalence between steps (A7) and (V7) is guaranteed by the discussion in [20]. Since the phase error rate $e^{(\mathrm{ph},\nu)}$ cannot be obtained directly in the actual protocol, we need to estimate $e^{(\mathrm{ph},\nu)}$ with some statistics such as the disturbance information during Alice and Bob's quantum communication.

C. POVM elements for the bit and phase errors

In this subsection, we construct POVMs for the bit and phase errors to estimate the upper bound on the phase error rate. To derive the relation between the bit and the phase error rates, we consider a measurement on Alice and Bob's quantum registers A and B just after the event is *detected* at step (V2), and regard an outcome as the occurrence of a bit error or a phase error. In this subsection, we explain only the definition and the resulting forms of POVMs for the bit and phase errors. The detailed derivations are referred to Appendix B.

The POVM element corresponding to the bit error in the pair of pulses (i,j) for $i<j$ is defined as

$$\begin{aligned}\hat{e}^{(\mathrm{b})}_{i,j} :=& \sum_{s_i,s_j\in\{0,1\}} \hat{P}(\hat{H}\,|s_i\rangle_{\mathrm{A},i})\hat{P}(\hat{H}\,|s_j\rangle_{\mathrm{A},j}) \\ &\otimes \frac{1}{|\mathcal{R}|}\sum_{r\in\mathcal{R}}\left[\hat{E}^{(r)}_{i,s_i\oplus s_j\oplus 1}\delta_{j,i+r} + \hat{E}^{(r)}_{j,s_i\oplus s_j\oplus 1}\delta_{i,j+r-L}\right] \\ =& \frac{2}{|\mathcal{R}|}\sum_{s_i,s_j\in\{0,1\}} \hat{P}(\hat{H}\,|s_i\rangle_{\mathrm{A},i})\hat{P}(\hat{H}\,|s_j\rangle_{\mathrm{A},j})\otimes \hat{E}^{(j-i)}_{i,s_i\oplus s_j\oplus 1}.\end{aligned} \tag{15}$$

Here, $|\mathcal{R}|$ denotes the cardinality of the set $\mathcal{R}$, and we use Eq. (5) in the last equality. Note that we omit identity operators on the subsystems. Next, the POVM element for a phase error is defined by the instances where Alice and Bob measure their qubits with the Hadamard basis $\{\hat{H}\,|0\rangle\,,\hat{H}\,|1\rangle\}$ and their outcomes disagree. The POVM element corresponding to the occurrence of the phase error in the pair of pulses (i,j) for $i<j$ is given by

$$\begin{aligned}\hat{e}^{(\mathrm{ph})}_{i,j} :=& \sum_{s=0}^{1}\sum_{k=1}^{3}\hat{M}^{(i,j)\dagger}_k\hat{P}(\hat{H}\,|s\rangle_{\mathrm{Aq}})\hat{M}^{(i,j)}_k \\ &\otimes\frac{1}{|\mathcal{R}|}\sum_{r\in\mathcal{R}}\left[\hat{F}^{(r)\dagger}_i\hat{P}(\hat{H}\,|\bar{s}\rangle_{\mathrm{B}})\hat{F}^{(r)}_i\delta_{j,i+r}\right. \\ &\left.+\hat{F}^{(r)\dagger}_j\hat{P}(\hat{H}\,|\bar{s}\rangle_{\mathrm{B}})\hat{F}^{(r)}_j\delta_{i,j+r-L}\right] \\ =& \frac{1}{|\mathcal{R}|}[\hat{P}(|0\rangle_{\mathrm{A},i}\,|1\rangle_{\mathrm{A},j})\otimes\hat{P}(|i\rangle_{\mathrm{B}}) \\ &+\hat{P}(|1\rangle_{\mathrm{A},i}\,|0\rangle_{\mathrm{A},j})\otimes\hat{P}(|j\rangle_{\mathrm{B}})] \\ &+\frac{1}{2|\mathcal{R}|}\sum_{t=0}^{1}\hat{P}(|t\rangle_{\mathrm{A},i}\,|t\rangle_{\mathrm{A},j})\otimes[\hat{P}(|i\rangle_{\mathrm{B}})+\hat{P}(|j\rangle_{\mathrm{B}})],\end{aligned} \tag{16}$$

where $\bar{s}=s\oplus 1$. For simplicity of analysis, we introduce a unitary operator $\hat{U}$ acting on $\mathcal{H}_{\mathrm{A}}\otimes\mathcal{H}_{\mathrm{B}}$ defined by

$$\hat{U}\bigotimes_{k'=1}^{L}(\hat{H}\,|s_{k'}\rangle_{\mathrm{A},k'})\,|k\rangle_{\mathrm{B}} = (-1)^{s_k}\bigotimes_{k'=1}^{L}(\hat{H}\,|s_{k'}\rangle_{\mathrm{A},k'})\,|k\rangle_{\mathrm{B}}\,. \tag{17}$$

By using $\hat{U}$ and Eq. (15), it is straightforward to show that

$$\hat{U}\hat{e}^{(\mathrm{b})}_{i,j}\hat{U}^\dagger = \frac{2}{|\mathcal{R}|}\hat{1}_{\mathrm{A}}\otimes\hat{E}^{(j-i)}_{i,1} = \hat{1}_{\mathrm{A}}\otimes\frac{1}{2|\mathcal{R}|}\hat{P}(|i\rangle_{\mathrm{B}}-|j\rangle_{\mathrm{B}}). \tag{18}$$

Since $\hat{U}$ also satisfies

$$\hat{U}\hat{P}(|s\rangle_{\mathrm{A},k})\hat{P}(|k'\rangle_{\mathrm{B}})\hat{U}^\dagger = \hat{P}(|s\oplus\delta_{k,k'}\rangle_{\mathrm{A},k})\hat{P}(|k'\rangle_{\mathrm{B}}), \tag{19}$$

we have

$$\begin{aligned}\hat{U}\hat{e}^{(\mathrm{ph})}_{i,j}\hat{U}^\dagger =& \frac{1}{2|\mathcal{R}|}[\hat{P}(|1\rangle_{\mathrm{A},i})+\hat{P}(|1\rangle_{\mathrm{A},j})] \\ &\otimes[\hat{P}(|i\rangle_{\mathrm{B}})+\hat{P}(|j\rangle_{\mathrm{B}})].\end{aligned} \tag{20}$$

For the state $\hat{\rho}$ of a *detected* event, the probability of having a bit error in the extracted qubit pair Aq and Bq is expressed by $\mathrm{Tr}(\hat{\rho}\hat{e}^{(\mathrm{b})})$, while a phase error is given by $\mathrm{Tr}(\hat{\rho}\hat{e}^{(\mathrm{ph})})$, where $\hat{e}^{(\mathrm{b})}$ and $\hat{e}^{(\mathrm{ph})}$ are, respectively, given by

$$\hat{e}^{(\mathrm{b})} = \sum_{(i,j):j-i\in\mathcal{R}}\hat{e}^{(\mathrm{b})}_{i,j},\quad \hat{e}^{(\mathrm{ph})} = \sum_{(i,j):j-i\in\mathcal{R}}\hat{e}^{(\mathrm{ph})}_{i,j}. \tag{21}$$

By applying $\hat{U}$ to $\hat{e}^{(\mathrm{b})}$ in Eq. (21) and using Eq. (18), we have

$$\hat{U}\hat{e}^{(\mathrm{b})}\hat{U}^\dagger = \hat{1}_{\mathrm{A}}\otimes\hat{\Pi}^{(\mathrm{b})}, \tag{22}$$

where the matrix elements of $\hat{\Pi}^{(\mathrm{b})}$ are

$${}_{\mathrm{B}}\langle m|\,\hat{\Pi}^{(\mathrm{b})}\,|n\rangle_{\mathrm{B}} = \begin{cases}\frac{1}{2} & \text{if } m=n,\\ -\frac{1}{2|\mathcal{R}|} & \text{if } |m-n|\in\mathcal{R},\\ 0 & \text{otherwise.}\end{cases} \tag{23}$$

By applying $\hat{U}$ to $\hat{e}^{(\mathrm{ph})}$, using Eq. (20) and $\hat{P}(|k\rangle_{\mathrm{A},m}) = \sum_{\vec{a}}\hat{P}(|\vec{a}\rangle_{\mathrm{A}})\delta_{a_m,k}$ with $\vec{a}=a_1a_2\ldots a_L$ and $|\vec{a}\rangle_{\mathrm{A}} = |a_1\rangle_{\mathrm{A},1}\,|a_2\rangle_{\mathrm{A},2}\ldots|a_L\rangle_{\mathrm{A},L}$, $\hat{U}\hat{e}^{(\mathrm{ph})}\hat{U}^\dagger$ results in the following form.

$$\begin{aligned}\hat{U}\hat{e}^{(\mathrm{ph})}\hat{U}^\dagger =& \sum_{\vec{a}}\hat{P}(|\vec{a}\rangle_{\mathrm{A}}) \\ &\otimes\sum_{m=1}^{L}\hat{P}(|m\rangle_{\mathrm{B}})\left(\frac{1}{2}\delta_{a_m,1}+\frac{1}{2|\mathcal{R}|}\sum_{n:|m-n|\in\mathcal{R}}\delta_{a_n,1}\right) \\ =:& \sum_{\vec{a}}\hat{P}(|\vec{a}\rangle_{\mathrm{A}})\otimes\hat{\Pi}^{(\mathrm{ph})}_{\vec{a}}.\end{aligned} \tag{24}$$

D. Relations between the bit and the phase error rates

In this subsection, we derive the upper bound on $h^{(\mathrm{ph})}$ in Eq. (13) by using the bit error rate. For this, we first derive the range where Alice's sending state can be contained. In the virtual protocol, if the initial state $|\Phi\rangle$ satisfies $\langle\Phi|\,|\vec{a}\rangle_{\mathrm{A}}\,|\nu\rangle_{\mathrm{C}} = 0$ for a state $|\vec{a}\rangle_{\mathrm{A}}\,|\nu\rangle_{\mathrm{C}}$, the density operator $\hat{\rho}$ of Aq and Bq originating from $|\Phi\rangle$ also satisfies $\hat{\rho}\,|\vec{a}\rangle_{\mathrm{A}}\,|\nu\rangle_{\mathrm{C}} = 0$. Moreover, we have the following relations between $\vec{a}$ and ν such that $\langle\Phi|\,|\vec{a}\rangle_{\mathrm{A}}\,|\nu\rangle_{\mathrm{C}} = 0$ is satisfied [9],

$${}_{\mathrm{A}}\langle\vec{a}|_{\mathrm{C}}\,\langle\nu|\,|\Phi\rangle = 0 \ \text{ if } |\vec{a}|>\nu, \tag{25}$$

$${}_{\mathrm{A}}\langle\vec{a}|_{\mathrm{C}}\,\langle\nu|\,|\Phi\rangle = 0 \ \text{ if } (-1)^{|\vec{a}|}\neq(-1)^{\nu}, \tag{26}$$

where $|\vec{a}|$ denotes the number of 1's in the bit string $\vec{a}$. By using Eqs. (25) and (26), $\hat{\rho}$ after Alice obtains ν by measuring the system C is contained in the range of a projection operator $\hat{P}^{(\nu)}$, which is defined by

$$\hat{P}^{(\nu)} := \sum_{\vec{a}:|\vec{a}|=\nu,\nu-2,\nu-4,\ldots}\sum_{i=1}^{L}\hat{P}(|\vec{a}\rangle_{\mathrm{A}}\otimes|i\rangle_{\mathrm{B}}). \tag{27}$$

Next, to derive the relation between the bit and phase error rates, we consider the quantity $\Omega^{(\nu)}(\lambda)$ defined as the largest eigenvalue of the operator,

$$\hat{P}^{(\nu)}(\hat{e}^{(\mathrm{ph})}-\lambda\hat{e}^{(\mathrm{b})})\hat{P}^{(\nu)}, \tag{28}$$

in the range of $\hat{P}^{(\nu)}$ with $\lambda\geqslant 0$. By using $\Omega^{(\nu)}(\lambda)$, the phase error rate $e^{(\mathrm{ph},\nu)}$ when Alice emits ν photons is bounded by the bit error rate $e^{(\mathrm{b},\nu)}$ when Alice emits ν photons as [9]

$$e^{(\mathrm{ph},\nu)}\leqslant\lambda e^{(\mathrm{b},\nu)}+\Omega^{(\nu)}(\lambda). \tag{29}$$

Since Eq. (29) for various $\lambda \geqslant 0$ determines a convex achievable region of $(e^{(\mathrm{b},\nu)}, e^{(\mathrm{ph},\nu)})$ and $h(x)$ is monotonically increasing, we obtain the convex achievable region of $(e^{(\mathrm{b},\nu)}, h(e^{(\mathrm{ph},\nu)}))$ specified by

$$h(e^{(\mathrm{ph},\nu)}) \leqslant \gamma e^{(\mathrm{b},\nu)} + \Omega_h^{(\nu)}(\gamma) \tag{30}$$

for various $\gamma \geqslant 0$. Here, $\Omega_h^{(\nu)}(\gamma)$ is the quantity depending on γ and $\Omega^{(\nu)}(\lambda)$.

In order to derive an upper bound on the leaked information $h^{(\mathrm{ph})}$ in Eq. (14), we consider the optimization of $Q^{(\nu)}$ such that Eve's information is maximal. Since $NQ^{(\nu)}$ is the number of qubits extracted in step (V5) from the *detected* events when Alice emits ν photons, $Q^{(\nu)}$ needs to satisfy the following physical requirement regarding the number of total events when Alice emits ν photons,

$$NQ^{(\nu)} \leqslant Np_\nu, \tag{31}$$

where p_ν denotes the Poisson distribution with mean $L\alpha^2$,

$$p_\nu := e^{-L\alpha^2} \frac{(L\alpha^2)^\nu}{\nu!}. \tag{32}$$

Let $q^{(\nu)}$ be a fraction of *detected* events when Alice emits ν photons among all the detection,

$$q^{(\nu)} = \frac{Q^{(\nu)}}{Q} = \frac{Q^{(\nu)}}{\sum_{\nu=0}^{\infty} Q^{(\nu)}}. \tag{33}$$

Here, $q^{(\nu)}$ is chosen by Eve under the constraint of Eq. (31). As long as $\Omega^{(0)}(\lambda) \leqslant \Omega^{(1)}(\lambda) \leqslant \Omega^{(2)}(\lambda) \leqslant \ldots$ holds for all $\lambda \geqslant 0$, Eve can maximize the amount of leaked information by using the events with a larger value of ν. Therefore, the optimal strategy for Eve is the following choice:

$$q^{(\nu)} = \begin{cases} Q^{-1} p_\nu & \text{if} \quad \nu \geqslant \nu_0 + 1, \\ 1 - Q^{-1}\left(1 - \sum_{\nu'=0}^{\nu_0} p_{\nu'}\right) & \text{if} \quad \nu = \nu_0, \\ 0 & \text{if} \quad \nu \leqslant \nu_0 - 1, \end{cases} \tag{34}$$

where ν_0 is the integer satisfying

$$1 - \sum_{\nu'=0}^{\nu_0} p_{\nu'} < Q \leqslant 1 - \sum_{\nu'=0}^{\nu_0 - 1} p_{\nu'}. \tag{35}$$

By using $\{q^{(\nu)}\}_\nu$ and Eq. (30), the upper bound on $h^{(\mathrm{ph})}$ in Eq. (14) is written as

$$h^{(\mathrm{ph})} = \sum_{\nu=0}^{\infty} q^{(\nu)} h(e^{(\mathrm{ph},\nu)}) \leqslant \min_{\gamma \geqslant 0} \left\{ \gamma e^{(\mathrm{b})} + \sum_{\nu=0}^{\infty} q^{(\nu)} \Omega_h^{(\nu)}(\gamma) \right\}, \tag{36}$$

where $e^{(\mathrm{b})} = \sum_\nu q^{(\nu)} e^{(\mathrm{b},\nu)}$ is the bit error rate in the actual protocol. The task left to obtain the upper bound on $h^{(\mathrm{ph})}$ is to evaluate the quantities $\Omega^{(\nu)}(\lambda)$ for $\nu \in [0,\infty)$. In our analysis, we consider the upper bounds on $\Omega^{(\nu)}(\lambda)$ for $\nu = 1,2$, and for $\nu \geqslant 3$ we make a pessimistic assumption that all the information is leaked to Eve, that is, $\Omega^{(\nu)}(\lambda) \equiv 1$. With this consideration, Eq. (36) is upper bounded by

$$h^{(\mathrm{ph})} \leqslant \min_{\gamma \geqslant 0} \left\{ \gamma e^{(\mathrm{b})} + \sum_{\nu=0}^{2} q^{(\nu)} \Omega_h^{(\nu)}(\gamma) + q^{(\nu \geqslant 3)} \right\}, \tag{37}$$

with $q^{(\nu \geqslant 3)} := \sum_{\nu=3}^{\infty} q^{(\nu)}$.

E. Evaluation of $\Omega^{(\nu)}(\lambda)$

To evaluate $\Omega^{(\nu)}(\lambda)$, we apply the unitary operation $\hat{U}$ in Eq. (17) to Eq. (28), and we obtain [9]

$$\begin{aligned} &\hat{U}\hat{P}^{(\nu)}\hat{U}^\dagger(\hat{U}\hat{e}^{(\mathrm{ph})}\hat{U}^\dagger - \lambda \hat{U}\hat{e}^{(\mathrm{b})}\hat{U}^\dagger)\hat{U}\hat{P}^{(\nu)}\hat{U}^\dagger \\ &= \sum_{\vec{a}:|\vec{a}|=\nu-1,\nu-3,\ldots} \hat{P}(|\vec{a}\rangle_{\mathrm{A}}) \otimes \left(\hat{\Pi}_{\vec{a}}^{(\mathrm{ph})} - \lambda \hat{\Pi}^{(\mathrm{b})}\right) \\ &\quad + \sum_{\vec{a}:|\vec{a}|=\nu+1} \hat{P}(|\vec{a}\rangle_{\mathrm{A}}) \otimes \hat{P}_{\vec{a}} \left(\hat{\Pi}_{\vec{a}}^{(\mathrm{ph})} - \lambda \hat{\Pi}^{(\mathrm{b})}\right) \hat{P}_{\vec{a}}. \end{aligned} \tag{38}$$

Here, we use the following equation:

$$\begin{aligned} \hat{U}\hat{P}^{(\nu)}\hat{U}^\dagger &= \sum_{\vec{a}:|\vec{a}|=\nu-1,\nu-3,\ldots} \hat{P}(|\vec{a}\rangle_{\mathrm{A}}) \otimes \hat{1}_{\mathrm{B}} \\ &\quad + \sum_{\vec{a}:|\vec{a}|=\nu+1} \hat{P}(|\vec{a}\rangle_{\mathrm{A}}) \otimes \hat{P}_{\vec{a}}, \end{aligned} \tag{39}$$

where

$$\hat{P}_{\vec{a}} := \sum_{i=1}^{L} \delta_{a_i,1} \hat{P}(|i\rangle_{\mathrm{B}}). \tag{40}$$

Since $\hat{\Pi}_{\vec{a}}^{(\mathrm{ph})} \geqslant \hat{\Pi}_{\vec{a}'}^{(\mathrm{ph})}$ for $(\vec{a},\vec{a}')$ such that $a_i \geqslant a_i'$ for all i, we can neglect the operators with $\vec{a}$ satisfying $|\vec{a}| \leqslant \nu - 3$. Therefore, we have only to consider $\Omega_-^{(\nu)}(\lambda)$ defined as the largest eigenvalue of the operators,

$$\left\{\hat{\Pi}_{\vec{a}}^{(\mathrm{ph})} - \lambda \hat{\Pi}^{(\mathrm{b})} \mid |\vec{a}| = \nu - 1\right\}, \tag{41}$$

and $\Omega_+^{(\nu)}(\lambda)$ defined as the largest eigenvalue of the operators,

$$\left\{\hat{P}_{\vec{a}}\left(\hat{\Pi}_{\vec{a}}^{(\mathrm{ph})} - \lambda \hat{\Pi}^{(\mathrm{b})}\right)\hat{P}_{\vec{a}} \mid |\vec{a}| = \nu + 1\right\}. \tag{42}$$

Let $e_\pm^{(\mathrm{ph},\nu)}$ be the upper bound on the phase error rate when we employ $\Omega_\pm^{(\nu)}(\lambda)$, respectively. $e_\pm^{(\mathrm{ph},\nu)}$ is given by

$$e_\pm^{(\mathrm{ph},\nu)} = \min_{\lambda \geqslant 0} \{\lambda e^{(\mathrm{b},\nu)} + \Omega_\pm^{(\nu)}(\lambda)\}. \tag{43}$$

Regarding $\Omega_+^{(\nu)}(\lambda)$, we have an analytical formula that is given in the following lemma (see Appendix C for its proof).

Lemma 2. If $\nu \leqslant |\mathcal{R}|/2$,

$$\Omega_+^{(\nu)}(\lambda) = \frac{1-\lambda}{2} + \nu \frac{1+\lambda}{2|\mathcal{R}|} \tag{44}$$

holds for arbitrary $\lambda \geqslant 0$.

By applying Lemma 2 to Eq. (43), we obtain the analytical solution for the upper bound on $e_+^{(\mathrm{ph},\nu)}$.

$$e_+^{(\mathrm{ph},\nu)} \leqslant \min_{\lambda\geqslant 0}\left\{\lambda e^{(\mathrm{b},\nu)} + \frac{1-\lambda}{2} + \nu\frac{1+\lambda}{2|\mathcal{R}|}\right\} = \frac{|\mathcal{R}|+\nu}{2|\mathcal{R}|} + \min_{\lambda\geqslant 0}\left\{\lambda\left(e^{(\mathrm{b},\nu)} - \frac{|\mathcal{R}|-\nu}{2|\mathcal{R}|}\right)\right\}\begin{cases}\leqslant 0 & \text{if } e^{(\mathrm{b},\nu)} < \frac{|\mathcal{R}|-\nu}{2|\mathcal{R}|} = \frac{1}{2} - \frac{\nu}{2|\mathcal{R}|},\\ = \frac{|\mathcal{R}|+\nu}{2|\mathcal{R}|} = \frac{1}{2} + \frac{\nu}{2|\mathcal{R}|} & \\ & \text{if } e^{(\mathrm{b},\nu)} \geqslant \frac{|\mathcal{R}|-\nu}{2|\mathcal{R}|} = \frac{1}{2} - \frac{\nu}{2|\mathcal{R}|},\end{cases} \tag{45}$$

where Eq. (45) is independent of the length L of one block. Note that if $\nu > |\mathcal{R}|/2$, Lemma 2 cannot be applied, however, $\Omega_+^{(\nu)}(\lambda)$ for this case is also easily derived by following the same discussion of the derivation of Eq. (44) (the discussion in Appendix C covers this situation, for example, $|\mathcal{R}| = \nu = 2$). On the other hand, to derive the upper bound $\Omega_-^{(\nu)}(\lambda)$ in Eq. (41), we need to solve the largest eigenvalue of $\hat{\Pi}_{\vec{a}}^{(\mathrm{ph})} - \lambda\hat{\Pi}^{(\mathrm{b})}$, which is written as

$$\begin{aligned}&\hat{\Pi}_{\vec{a}}^{(\mathrm{ph})} - \lambda\hat{\Pi}^{(\mathrm{b})}\\ &= \sum_{m=1}^{L}\hat{P}(|m\rangle_{\mathrm{B}})\left(\frac{1}{2}\delta_{a_m,1} + \frac{1}{2|\mathcal{R}|}\sum_{n:|m-n|\in\mathcal{R}}\delta_{a_n,1}\right)\\ &\quad - \lambda\left(\frac{1}{2}\hat{1}_{\mathrm{B}} - \frac{1}{2|\mathcal{R}|}\sum_{(m,n):|m-n|\in\mathcal{R}}|m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|\right)\\ &= \sum_{m=1}^{L}\hat{P}(|m\rangle_{\mathrm{B}})\left(\frac{\delta_{a_m,1}-\lambda}{2} + \frac{1}{2|\mathcal{R}|}\sum_{n:|m-n|\in\mathcal{R}}\delta_{a_n,1}\right)\\ &\quad + \frac{\lambda}{2|\mathcal{R}|}\sum_{(m,n):|m-n|\in\mathcal{R}}|m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|\,.\end{aligned} \tag{46}$$

To discuss the eigenvalues of Eq. (46), we recall the fact that the translation operation $\hat{V}_{\vec{a},\vec{a}'}$ defined for $(\vec{a},\vec{a}')$ such that $a_k = a'_{k+_L\kappa}$ is satisfied for any $k \in \{1,\ldots,L\}$ with specific $\kappa \in \{1,\ldots,L\}$,

$$\hat{V}_{\vec{a},\vec{a}'} := \sum_{m=1}^{L}|m +_L \kappa\rangle_{\mathrm{B}\,\mathrm{B}}\langle m| \tag{47}$$

does not change the eigenvalues. Hence, the eigenvalues of Eq. (46) with $\vec{a}$ and $\vec{a}'$ are the same if there exists κ ($1 \leqslant \kappa \leqslant L$) such that $a_k = a'_{k+_L\kappa}$ is satisfied for any $k \in \{1,\ldots,L\}$. By using this, for $\nu = 2$, it is enough to consider the case $\vec{a} = (\overbrace{0\ldots0}^{|\mathcal{R}|/2}10\ldots0)$, where the matrix representation of Eq. (46) is written as

$$\begin{aligned}&\frac{1}{2|\mathcal{R}|}\mathrm{diag}\{\overbrace{1,\ldots,1}^{|\mathcal{R}|/2},|\mathcal{R}|,\overbrace{1,\ldots,1}^{|\mathcal{R}|/2},0\ldots0\}_L\\ &\quad - \frac{\lambda}{2}I_L + \frac{\lambda}{2|\mathcal{R}|}A_L^{(|\mathcal{R}|/2)}.\end{aligned} \tag{48}$$

Here, $A_n^{(k)}$ is an $n\times n$ matrix satisfying

$$\left(A_n^{(k)}\right)_{l,m} = \begin{cases}1 & \text{if} \quad 1 \leqslant |l-m| \leqslant k,\\ 1 & \text{if} \quad n-k \leqslant |l-m| \leqslant n-1,\\ 0 & \text{otherwise.}\end{cases} \tag{49}$$

In order to obtain the largest eigenvalue of Eq. (48) for large L, we take a numerical approach. For $\nu = 1$, however, the upper bound on $\Omega_-^{(\nu)}(\lambda)$ can be easily derived, and we obtain the following theorem for the upper bound on $e^{(\mathrm{ph},1)}$ (see Appendix D for its proof).

Theorem 1. For $0 \leqslant e^{(\mathrm{b},1)} \leqslant \frac{|\mathcal{R}|-1}{2|\mathcal{R}|}$,

$$e^{(\mathrm{ph},1)} \leqslant \frac{|\mathcal{R}|+1}{|\mathcal{R}|-1}e^{(\mathrm{b},1)} \tag{50}$$

holds. Also, for $e^{(\mathrm{b},1)} \geqslant \frac{|\mathcal{R}|-1}{2|\mathcal{R}|}$, we have

$$e^{(\mathrm{ph},1)} \leqslant \frac{1}{2} + \frac{1}{2|\mathcal{R}|}. \tag{51}$$

For $\nu > 1$, the upper bound on $e^{(\mathrm{ph},\nu)}$ is derived as the maximum value of the convex combination of the one on $e_+^{(\mathrm{ph},\nu)}$ and $e_-^{(\mathrm{ph},\nu)}$ as

$$e^{(\mathrm{ph},\nu)} \leqslant \max_{\substack{p\,:\,0\leqslant p\leqslant 1\\ px_+ + (1-p)x_- = e^{(\mathrm{b},\nu)}}}\{pf_+^{(\nu)}(x_+) + (1-p)f_-^{(\nu)}(x_-)\}, \tag{52}$$

where $f_\pm^{(\nu)}(x)$ is defined by

$$f_\pm^{(\nu)}(x) = \min_{\lambda_\pm\geqslant 0}\{\lambda_\pm x + \Omega_\pm^{(\nu)}(\lambda_\pm)\}. \tag{53}$$

In Figs. 4(a) and 4(b), we plot the estimated upper bound on the phase error rate $e^{(\mathrm{ph},\nu)}$ as a function of $e^{(\mathrm{b},\nu)}$ with the number of delays $|\mathcal{R}| \in \{2,4,6,8,10\}$ for $\nu \in \{1,2\}$. From these figures, even if the number of the random delays is small (say $|\mathcal{R}| = 2$), we can observe a significant improvement over the original DPS protocol. Moreover, when we compare the resulting phase error rate of the SNRDPS protocol (solid lines) with the one of the RRDPS protocol (dashed lines) with the small random delays, the phase error rate of the SNRDPS protocol is smaller if the bit error rate is small. Here, we assume the phase error rate of the RRDPS protocol as $e^{(\mathrm{ph})} = \nu/|\mathcal{R}|$ [11].

IV. KEY GENERATION RATES

In this section, we present our main results, namely, the key generation rate of the SNRDPS protocol is significantly enhanced over the one of the original DPS protocol only by employing a few additional delays such as $|\mathcal{R}| = 2$. In Fig. 5, we compare the key generation rate G per pulse in Eq. (13) for three protocols: (i) the original DPS protocol [9], (ii) the RRDPS protocol [11] without monitoring the disturbance when Bob employs the number of random delays $|\mathcal{R}|$, and (iii) the SNRDPS protocol when Bob employs the number of random delays $|\mathcal{R}|$. From Fig. 5(a), we can see that the key generation rate of the SNRDPS protocol with $|\mathcal{R}| = 2$

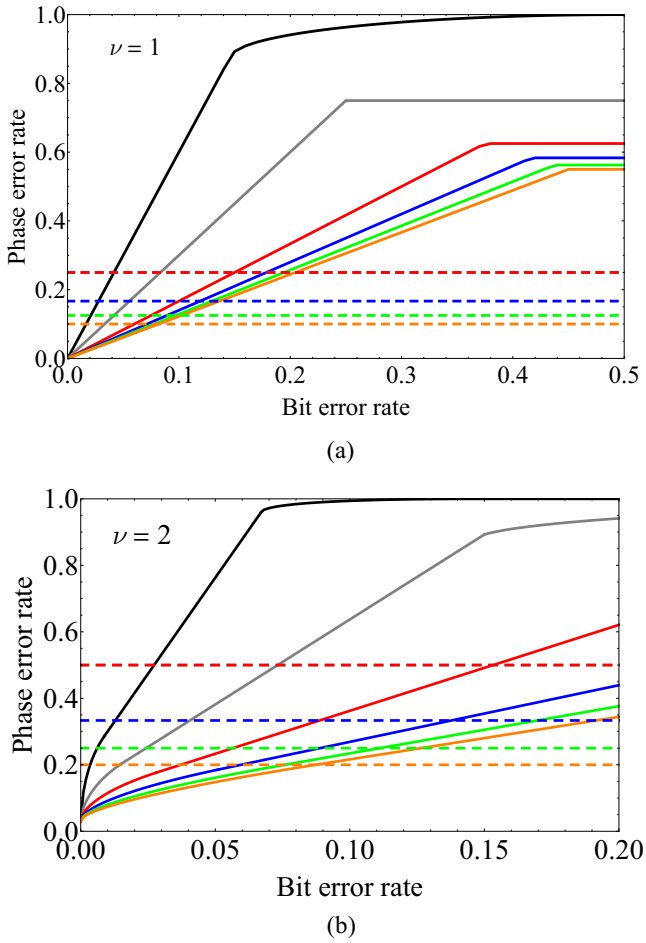


FIG. 4. The upper bound on the phase error rate $e^{(\mathrm{ph},\nu)}$ as a function of the bit error rate $e^{(\mathrm{b},\nu)}$ in the case of (a) $\nu = 1$ and (b) $\nu = 2$, respectively. From bottom to top, the solid lines correspond to the case of the SNRDPS protocol with $(L,|\mathcal{R}|) = (32,10),(32,8),(32,6),(32,4)$, and $(32,2)$, respectively. The black solid line on the topmost of both figures corresponds to the case for the original DPS protocol [9] with $L = 32$. Also, from bottom to top, the dashed lines correspond to the case of the RRDPS protocol [11] with $(L,|\mathcal{R}|) = (11,10),(9,8),(7,6)$, and $(5,4)$, respectively.

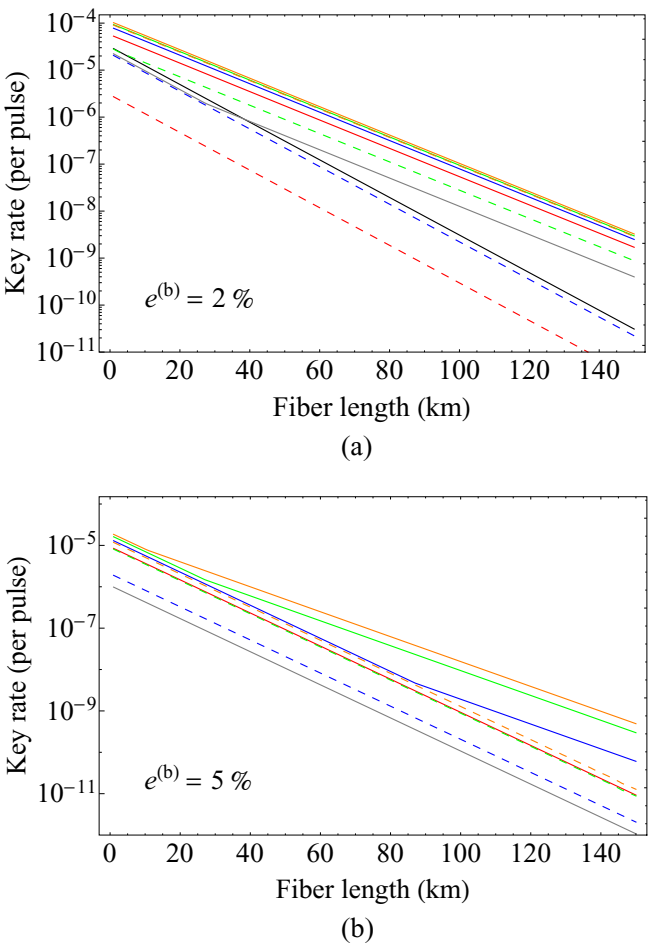


FIG. 5. The key generation rate G in logarithmic scale vs fiber length. We assume Bob's detection efficiency of 10%, the channel transmittance of 0.2 dB/km, and the bit error rate $e^{(\mathrm{b})} = 2\%$ in (a) and $e^{(\mathrm{b})} = 5\%$ in (b), respectively. In this simulation, we have optimized the mean photon number $\mu = |\alpha|^2$. From top to bottom, the solid lines are for the SNRDPS protocols with $(L,|\mathcal{R}|) = (32,10),(32,8),(32,6),(32,4)$, and $(32,2)$, respectively. The key generation rate of the original DPS protocol [9] with $L = 32$ is plotted in the bottommost solid line in (a). In the original DPS protocol, the secret key cannot be extracted when $e^{(\mathrm{b})} = 5\%$. Also, from top to bottom, the dashed lines express the resulting key generation rates for the RRDPS protocol with $(L,|\mathcal{R}|) = (11,10),(9,8),(7,6)$, and $(5,4)$, respectively. Note that the dashed lines corresponding to $(L,|\mathcal{R}|) = (5,4)$ disappeared from (b) since the rate is zero.

outperforms the original DPS protocol when the fiber length is more than about 40 km and the bit error rate $e^{(\mathrm{b})}$ is 2%. Also, the SNRDPS protocol always outperforms the original DPS protocol when $|\mathcal{R}| \geqslant 4$ and $e^{(\mathrm{b})} = 2\%$. Moreover, from Fig. 5(b), the SNRDPS protocol provides a positive key generation rate even though the original DPS protocol cannot generate the secret key. Also, in both figures, by comparing two lines with the same colors, we confirm that the SNRDPS protocol outperforms the RRDPS protocol with the same $|\mathcal{R}|$ up to $|\mathcal{R}| = 10$. This means that, if the amount of randomness is small as $2 \leqslant |\mathcal{R}| \leqslant 10$ and $e^{(\mathrm{b})} \leqslant 5\%$, the key generation rate of the SNRDPS protocol outperforms the one of the RRDPS protocol when both protocols employ the same number of random delays $|\mathcal{R}|$ and Alice and Bob do not monitor the bit error rate in the RRDPS protocol.

Next, we discuss the transmittance dependency of the key generation rates for the SNRDPS protocol. For this, we assume a fiber-based QKD system, and the detection efficiency Q is assumed to be $Q = \frac{L\mu\eta}{2} e^{-L\mu\eta}$. Here, $\mu := |\alpha|^2$ denotes the mean photon number per sending pulse, and $\eta = \eta(l)$ denotes channel transmittance with the fiber length l as $\eta(l) := \eta_0 \times 10^{-\frac{0.2l}{10}}$ with η_0 denoting Bob's detection efficiency. If μ is small ($\mu < 1$), Q and p_ν are approximated to $Q \sim O(\mu\eta)$ and $p_\nu \sim O(\mu^\nu)$, respectively. Suppose that Alice and Bob generate the secret key from the ν-photon emissions. In this case, by considering Eve's attack, the total detection efficiency $Q \sim O(\mu\eta)$ minus the probability of emitting more than $\nu + 1$ photons (approximated to $\mu^{\nu+1}$) has to be positive, resulting in $O(\mu^{\nu+1}) \leqslant O(\mu\eta)$. From this, we obtain the dependency of μ over the transmittance as

$$\mu \sim O(\eta^{1/\nu}), \tag{54}$$

and hence the key generation rate G behaves as

$$G \sim Q \sim O(\eta^{(\nu+1)/\nu}). \tag{55}$$

In Fig. 5(a), all the lines of the SNRDPS protocol except for the one with $(L,|\mathcal{R}|) = (32,2)$ and the one of the RRDPS protocol with $|\mathcal{R}| = 10$ result in the transmittance dependency as $G \sim O(\eta^{3/2})$, which means that the secret key is extracted from the two-photon emission events in addition to the single-photon emission events. On the other hand, the line of the original DPS protocol and the ones of the RRDPS protocol with $|\mathcal{R}| = 4,6$ result in the transmittance dependency as $G \sim O(\eta^2)$ since the upper bound on the phase error rate of the two-photon emission events is too high to extract the secret key. Moreover, the SNRDPS protocol with $(L,|\mathcal{R}|) = (32,2)$ and the RRDPS protocol with $|\mathcal{R}| = 8$ provide the key generation rate of the form $G \sim O(\eta^2)$ for shorter distance and $G \sim O(\eta^{3/2})$ for longer distance. The implication of this is that when the loss increases, the two-photon contribution becomes larger, and moreover the bit error rate of 2% is still small enough to generate the key from the two-photon emission event. Also, in Fig. 5(b), all the lines of the SNRDPS protocol except the one with $(L,|\mathcal{R}|) = (32,4)$ and $(32,2)$ result in the transmittance dependency as $G \sim O(\eta^2)$ for shorter distance and $G \sim O(\eta^{3/2})$ for longer distance, while all the remaining lines result in the transmittance dependency as $G \sim O(\eta^2)$. Even when the bit error rate is 5%, the properties of transmittance dependency of the key generation rates including the change of the scaling from the short and the long distance regimes can be explained with the same reason as the ones in Fig. 5(a).

Finally, Figs. 6(a) and 6(b) show the optimal mean photon number $L\mu$ to realize the key generation rates in Figs. 5(a) and 5(b), respectively. For all the protocols, it can be found that the optimal mean photon number scales as $O(\eta^{1/\nu})$ for the ν-photon emission event, which is the same scaling as Eq. (54). Also, the discontinuous point of the lines in Figs. 6(a) and 6(b), which represents the boundary of the presence or absence of the two-photon contribution, corresponds to the changing point of the scaling of the key generation rates in Figs. 5(a) and 5(b), respectively.

V. CONCLUSION

In conclusion, in this paper, we have proposed a DPS-type QKD protocol with small random delays at Bob's measurement and analyzed its information-theoretical security. For this protocol, we have estimated an upper bound on the phase error rate for Alice's single and two-photon emission events by using the bit error rate information. In addition, we have simulated and compared the key generation rates for the SNRDPS protocol with $|\mathcal{R}| \in \{2,4,6,8,10\}$, the one for the original DPS protocol, and the ones for the RRDPS protocol with $|\mathcal{R}| \in \{4,6,8,10\}$. As a result, we found that the performance of the SNRDPS protocol is significantly enhanced from the original DPS protocol even when Bob employs only a few numbers of delays such as $|\mathcal{R}| = 2$. Moreover, we found that if $|\mathcal{R}| \leqslant 10$, the key generation rate of the SNRDPS protocol based on our analysis outperforms the RRDPS protocol without monitoring the disturbance [11] when the same number of random delays is employed.

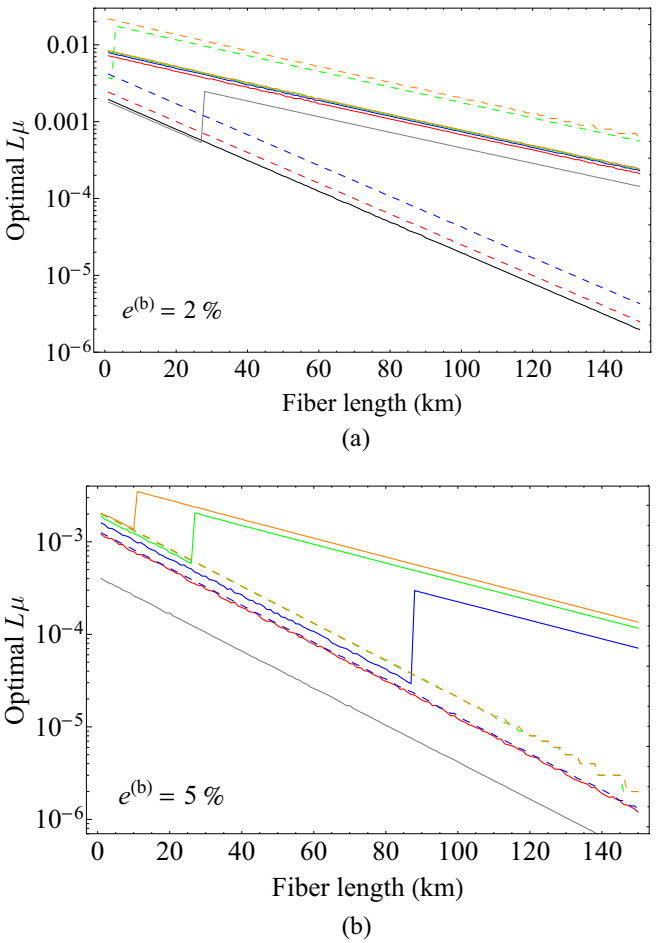


FIG. 6. The optimal light intensity $L\mu$ to achieve the key generation rate in Fig. 5 in logarithmic scale vs fiber length when the bit error rate $e^{(b)} = 2\%$ in (a) and $e^{(b)} = 5\%$ in (b). From top to bottom, the solid lines are for the cases of the SNRDPS protocol with $(L,|\mathcal{R}|) = (32,10),(32,8),(32,6),(32,4)$, and $(32,2)$, respectively. The light intensity of the original DPS protocol [9] with $L = 32$ is plotted in the bottom solid line in (a). In the original DPS protocol, the optimal light intensity cannot be found when $e^{(b)} = 5\%$. Also, from top to bottom, the dashed lines express the cases for the RRDPS protocol with $(L,|\mathcal{R}|) = (11,10),(9,8),(7,6)$, and $(5,4)$, respectively. Note that the dashed lines corresponding to $(L,|\mathcal{R}|) = (5,4)$ disappeared in (b) since the optimal light intensity cannot be found.

The SNRDPS protocol is an intermediate protocol between the original DPS and the RRDPS protocols in terms of the practicality and bit error tolerance, and this protocol increases the variety of future implementation for the DPS-type QKD protocol.

ACKNOWLEDGMENTS

We thank Toshihiko Sasaki, Hoi-Kwong Lo, Koji Azuma, Rikizo Ikuta, and Yuki Takeuchi for helpful discussions. N.I. acknowledges support from the MEXT/JSPS KAKENHI, Grant No. 16H02214. This work is in part funded by the ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan).

APPENDIX A: PROOF OF LEMMA 1

Here, we prove Lemma 1 in the main text. First, we have that $\hat{E}^{(r)}_{k,s}$ satisfies the following relation,

$$\hat{E}^{(r)}_{k,s} = \begin{cases} \hat{B}^{(r)}_{k,s} & \text{if } 1 \leqslant k \leqslant L-r, \\ \hat{B}^{(L-r)}_{k-L+r,s} & \text{if } L-r+1 \leqslant k \leqslant L. \end{cases} \tag{A1}$$

This is so because $\hat{E}^{(r)}_{k,s}$ is written as

$$\hat{E}^{(r)}_{k,s} = \frac{1}{2}\hat{P}\left(\frac{|k\rangle_{\rm B} + (-1)^s |k+r\rangle_{\rm B}}{\sqrt{2}}\right) = \hat{B}^{(r)}_{k,s} \tag{A2}$$

for $1 \leqslant k \leqslant L-r$ and

$$\begin{aligned}\hat{E}^{(r)}_{k,s} &= \frac{1}{2}\hat{P}\left(\frac{|k\rangle_{\rm B} + (-1)^s |k-L+r\rangle_{\rm B}}{\sqrt{2}}\right) \\ &= \frac{1}{2}\hat{P}\left(\frac{(-1)^s |k-L+r+(L-r)\rangle_{\rm B} + |k-L+r\rangle_{\rm B}}{\sqrt{2}}\right) \\ &= \hat{B}^{(L-r)}_{k-L+r,s}\end{aligned} \tag{A3}$$

for $L-r+1 \leqslant k \leqslant L$. By using Eq. (A1) and regarding $\hat{B}^{(r)}_{k,s} = 0$ if $k \leqslant 0$ or $L-r+1 \leqslant k$, we have the following equation:

$$\hat{E}^{(r)}_{k,s} = \hat{B}^{(r)}_{k,s} + \hat{B}^{(L-r)}_{k-L+r,s}. \tag{A4}$$

If we fix the delay of the dial measurement as r', the probability that Bob obtains the bit value s and announces the pair of integers (i,j) (with $i<j$) is given by

$$\begin{aligned}\Pr[s\wedge(i,j)|r']_{\rm dial} &= {\rm Tr}\big(\hat{\rho}\hat{E}^{(r')}_{i,s}\big)\delta_{j-i,r'} \\ &\quad + {\rm Tr}\big(\hat{\rho}\hat{E}^{(r')}_{j,s}\big)\delta_{j-i,L-r'}.\end{aligned} \tag{A5}$$

To simulate the dial measurement with the delay r' by using Bob's actual measurement, he randomly switches the delays of the actual measurement r' and $L-r'$. The probability that Bob obtains the outcome s and announces (i,j) when he performs the actual measurement is written as

$${\rm Tr}\big(\hat{\rho}\hat{B}^{(r')}_{i,s}\big)\delta_{j-i,r'} \tag{A6}$$

if the delay is r' and

$${\rm Tr}\big(\hat{\rho}\hat{B}^{(L-r')}_{i,s}\big)\delta_{j-i,L-r'} \tag{A7}$$

if the delay is $L-r'$. Let us define $\Pr[s\wedge(i,j)|r\in\{r',L-r'\}]_{\rm actual}$ as the probability that Bob obtains s and announces (i,j) when he performs the actual measurement with the delay $r=r'$ or $r=L-r'$ uniformly at random. $\Pr[s\wedge(i,j)|r\in\{r',L-r'\}]_{\rm actual}$ is written as

$$\begin{aligned}&\Pr[s\wedge(i,j)|r\in\{r',L-r'\}]_{\rm actual} \\ &= \tfrac{1}{2}{\rm Tr}\big(\hat{\rho}\hat{B}^{(r')}_{i,s}\big)\delta_{j-i,r'} + \tfrac{1}{2}{\rm Tr}\big(\hat{\rho}\hat{B}^{(L-r')}_{i,s}\big)\delta_{j-i,L-r'} \\ &= \tfrac{1}{2}{\rm Tr}\big(\hat{\rho}\hat{B}^{(r')}_{i,s}\big)\delta_{j-i,r'} + \tfrac{1}{2}{\rm Tr}\big(\hat{\rho}\hat{B}^{(L-r')}_{j-L+r',s}\big)\delta_{j-i,L-r'} \\ &= \tfrac{1}{2}{\rm Tr}\big(\hat{\rho}\big[\hat{E}^{(r')}_{i,s} - \hat{B}^{(L-r')}_{i-L+r',s}\big]\big)\delta_{j-i,r'} \\ &\quad + \tfrac{1}{2}{\rm Tr}\big(\hat{\rho}\big[\hat{E}^{(r')}_{j,s} - \hat{B}^{(r')}_{j,s}\big]\big)\delta_{j-i,L-r'} \\ &= \tfrac{1}{2}\big[{\rm Tr}\big(\hat{\rho}\hat{E}^{(r')}_{i,s}\big)\delta_{j-i,r'} + {\rm Tr}\big(\hat{\rho}\hat{E}^{(r')}_{j,s}\big)\delta_{j-i,L-r'}\big] \\ &\quad - \tfrac{1}{2}{\rm Tr}\big(\hat{\rho}\big[\hat{B}^{(L-r')}_{i-L+r',s}\delta_{j-i,r'} + \hat{B}^{(r')}_{j,s}\delta_{j-i,L-r'}\big]\big) \\ &= \tfrac{1}{2}\Pr[s\wedge(i,j)|r']_{\rm dial},\end{aligned} \tag{A8}$$

where we have used Eqs. (A4) and (A5) in the third and fifth equalities, respectively. Also, we have used

$$\begin{aligned}&\hat{B}^{(L-r')}_{i-L+r',s}\delta_{j-i,r'} + \hat{B}^{(r')}_{j,s}\delta_{j-i,L-r'} \\ &= \hat{B}^{(L-r')}_{i-L+r',s}\delta_{i-L+r',j-L} + \hat{B}^{(r')}_{j,s}\delta_{j,L-r'+i} \\ &= \hat{B}^{(L-r')}_{j-L,s}\delta_{i-L+r',j-L} + \hat{B}^{(r')}_{L-r'+i,s}\delta_{j,L-r'+i} = 0\end{aligned} \tag{A9}$$

in the fifth equality, which is satisfied since $\hat{B}^{(r)}_{k,s} = 0$ for $k \leqslant 0$ or $L-r+1 \leqslant k$, and (i,j) satisfies $1 \leqslant i < j \leqslant L$.

APPENDIX B: DETAIL OF CALCULATION OF BIT AND PHASE ERROR POVMs

Here, we detail the calculation of the equations of bit and phase error POVMs. First, Eq. (16) is derived as follows.

$$\begin{aligned}\hat{e}^{(\rm ph)}_{i,j} &:= \sum_{s=0}^{1}\sum_{k=1}^{3}\hat{M}^{(i,j)\dagger}_{k}\hat{P}(\hat{H}\,|s\rangle_{\rm Aq})\hat{M}^{(i,j)}_{k} \otimes \frac{1}{|\mathcal{R}|}\sum_{r\in\mathcal{R}}\big[\hat{F}^{(r)\dagger}_{i}\hat{P}(\hat{H}\,|\bar{s}\rangle_{\rm B})\hat{F}^{(r)}_{i}\delta_{j,i+r} + \hat{F}^{(r)\dagger}_{j}\hat{P}(\hat{H}\,|\bar{s}\rangle_{\rm B})\hat{F}^{(r)}_{j}\delta_{i,j+r-L}\big] \\ &= \frac{1}{|\mathcal{R}|}\sum_{s=0}^{1}\sum_{k=1}^{3}\hat{M}^{(i,j)\dagger}_{k}\hat{P}(\hat{H}\,|s\rangle_{\rm Aq})\hat{M}^{(i,j)}_{k} \otimes \big[\hat{F}^{(j-i)\dagger}_{i}\hat{P}(\hat{H}\,|\bar{s}\rangle_{\rm B})\hat{F}^{(j-i)}_{i} + \hat{F}^{(L-(j-i))\dagger}_{j}\hat{P}(\hat{H}\,|\bar{s}\rangle_{\rm B})\hat{F}^{(L-(j-i))}_{j}\big] \\ &= \frac{2}{|\mathcal{R}|}\sum_{s=0}^{1}\sum_{k=1}^{3}\hat{M}^{(i,j)\dagger}_{k}\hat{P}(\hat{H}\,|s\rangle_{\rm Aq})\hat{M}^{(i,j)}_{k} \otimes \hat{\mathcal{F}}^{\dagger}_{i,j}\hat{P}(\hat{H}\,|\bar{s}\rangle_{\rm B})\hat{\mathcal{F}}_{i,j} \\ &= \frac{1}{|\mathcal{R}|}\sum_{s=0}^{1}\left[\hat{P}(|s\rangle_{{\rm A},i}\,|\bar{s}\rangle_{{\rm A},j}) + \frac{1}{2}\sum_{t=0}^{1}\hat{P}(|t\rangle_{{\rm A},i}\,|t\rangle_{{\rm A},j})\right] \otimes \hat{P}(\delta_{\bar{s},1}\,|i\rangle_{\rm B} + \delta_{\bar{s},0}\,|j\rangle_{\rm B}) \\ &= \frac{1}{|\mathcal{R}|}[\hat{P}(|0\rangle_{{\rm A},i}\,|1\rangle_{{\rm A},j}) \otimes \hat{P}(|i\rangle_{\rm B}) + \hat{P}(|1\rangle_{{\rm A},i}\,|0\rangle_{{\rm A},j}) \otimes \hat{P}(|j\rangle_{\rm B})] + \frac{1}{2|\mathcal{R}|}\sum_{t=0}^{1}\hat{P}(|t\rangle_{{\rm A},i}\,|t\rangle_{{\rm A},j}) \otimes [\hat{P}(|i\rangle_{\rm B}) + \hat{P}(|j\rangle_{\rm B})].\end{aligned} \tag{B1}$$

Here, we define $\bar{s} := s\oplus 1$ and $\hat{\mathcal{F}}_{i,j} := \frac{1}{\sqrt{2}}\hat{H}\,|1\rangle_{{\rm Bq\,B}}\langle i| + \frac{1}{\sqrt{2}}\hat{H}\,|0\rangle_{{\rm Bq\,B}}\langle j|$.

Next, we detail the derivation of Eq. (22) as follows. $\hat{\Pi}^{(\mathrm{b})}$ in Eq. (22) is written as

$$\hat{\Pi}^{(\mathrm{b})} := \frac{1}{2|\mathcal{R}|}\sum_{(i,j):j-i\in\mathcal{R}} \hat{P}(|i\rangle_{\mathrm{B}} - |j\rangle_{\mathrm{B}}) = \frac{1}{2|\mathcal{R}|}\sum_{(i,j):j-i\in\mathcal{R}} [\hat{P}(|i\rangle_{\mathrm{B}}) + \hat{P}(|j\rangle_{\mathrm{B}}) - (|i\rangle_{\mathrm{B}\,\mathrm{B}}\langle j| + |j\rangle_{\mathrm{B}\,\mathrm{B}}\langle i|)]$$

$$= \frac{1}{2|\mathcal{R}|}\sum_{(m,n):|m-n|\in\mathcal{R}} [\hat{P}(|m\rangle_{\mathrm{B}}) - |m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|] = \frac{1}{2}\hat{1}_{\mathrm{B}} - \frac{1}{2|\mathcal{R}|}\sum_{(m,n):|m-n|\in\mathcal{R}} |m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|\,, \tag{B2}$$

which concludes Eq. (23).

Finally, we detail transformation in Eq. (24) as follows.

$$\hat{U}\hat{e}^{(\mathrm{ph})}\hat{U}^{\dagger} = \frac{1}{2|\mathcal{R}|}\sum_{(i,j):j-i\in\mathcal{R}} [\hat{P}(|1\rangle_{\mathrm{A},i}) + \hat{P}(|1\rangle_{\mathrm{A},j})] \otimes [\hat{P}(|i\rangle_{\mathrm{B}}) + \hat{P}(|j\rangle_{\mathrm{B}})] = \frac{1}{2|\mathcal{R}|}\sum_{(m,n):|m-n|\in\mathcal{R}} [\hat{P}(|1\rangle_{\mathrm{A},m}) + \hat{P}(|1\rangle_{\mathrm{A},n})] \otimes \hat{P}(|m\rangle_{\mathrm{B}})$$

$$= \frac{1}{2|\mathcal{R}|}\sum_{(m,n):|m-n|\in\mathcal{R}}\sum_{\vec{a}} \hat{P}(|\vec{a}\rangle_{\mathrm{A}})(\delta_{a_m,1} + \delta_{a_n,1}) \otimes \hat{P}(|m\rangle_{\mathrm{B}}) = \sum_{\vec{a}} \hat{P}(|\vec{a}\rangle_{\mathrm{A}}) \otimes \frac{1}{2|\mathcal{R}|}\sum_{m=1}^{L}\sum_{n:|m-n|\in\mathcal{R}} (\delta_{a_m,1} + \delta_{a_n,1})\hat{P}(|m\rangle_{\mathrm{B}})$$

$$= \sum_{\vec{a}} \hat{P}(|\vec{a}\rangle_{\mathrm{A}}) \otimes \sum_{m=1}^{L} \hat{P}(|m\rangle_{\mathrm{B}}) \left(\frac{1}{2}\delta_{a_m,1} + \frac{1}{2|\mathcal{R}|}\sum_{n:|m-n|\in\mathcal{R}} \delta_{a_n,1}\right) =: \sum_{\vec{a}} \hat{P}(|\vec{a}\rangle_{\mathrm{A}}) \otimes \hat{\Pi}^{(\mathrm{ph})}_{\vec{a}}. \tag{B3}$$

APPENDIX C: PROOF OF LEMMA 2

Here, we prove Lemma 2 in the main text. We consider the maximization of the largest eigenvalue of $\hat{P}_{\vec{a}}(\hat{\Pi}^{(\mathrm{ph})}_{\vec{a}} - \lambda\hat{\Pi}^{(\mathrm{b})})\hat{P}_{\vec{a}}$ in Eq. (42) over $\vec{a}$ with $|\vec{a}| = \nu + 1$. By using Eq. (46), $\hat{P}_{\vec{a}}(\hat{\Pi}^{(\mathrm{ph})}_{\vec{a}} - \lambda\hat{\Pi}^{(\mathrm{b})})\hat{P}_{\vec{a}}$ is written as

$$\hat{P}_{\vec{a}}\left(\hat{\Pi}^{(\mathrm{ph})}_{\vec{a}} - \lambda\hat{\Pi}^{(\mathrm{b})}\right)\hat{P}_{\vec{a}} = \sum_{m=1}^{L} \hat{P}(|m\rangle_{\mathrm{B}})\delta_{a_m,1}\left(\frac{\delta_{a_m,1} - \lambda}{2} + \frac{1}{2|\mathcal{R}|}\sum_{n:|m-n|\in\mathcal{R}} \delta_{a_n,1}\right) + \frac{\lambda}{2|\mathcal{R}|}\sum_{(m,n):|m-n|\in\mathcal{R}} |m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|\,\delta_{a_m,1}\delta_{a_n,1}$$

$$= \sum_{m:a_m=1}\left[\hat{P}(|m\rangle_{\mathrm{B}})\left(\frac{1-\lambda}{2} + \frac{1}{2|\mathcal{R}|}\sum_{n:|m-n|\in\mathcal{R}} \delta_{a_n,1}\right) + \frac{\lambda}{2|\mathcal{R}|}\sum_{n:|m-n|\in\mathcal{R}} |m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|\,\delta_{a_n,1}\right]$$

$$= \sum_{m:a_m=1}\left\{\frac{1-\lambda}{2}\hat{P}(|m\rangle_{\mathrm{B}}) + \sum_{n:|m-n|\in\mathcal{R}\wedge a_n=1}\left[\frac{1}{2|\mathcal{R}|}\hat{P}(|m\rangle_{\mathrm{B}}) + \frac{\lambda}{2|\mathcal{R}|}|m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|\right]\right\}$$

$$= \sum_{m:a_m=1}\frac{1-\lambda}{2}\hat{P}(|m\rangle_{\mathrm{B}}) + \sum_{\substack{(m,n):\\ |m-n|\in\mathcal{R}\wedge(a_m,a_n)=(1,1)}}\left[\frac{1}{2|\mathcal{R}|}\hat{P}(|m\rangle_{\mathrm{B}}) + \frac{\lambda}{2|\mathcal{R}|}|m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|\right]. \tag{C1}$$

Since $\mathcal{R} = \bigcup_{m=1}^{t}\{m, L-m\}$, the coefficient of $\hat{P}(|m\rangle_{\mathrm{B}})$ for m such that $a_m = 1$ in Eq. (C1) is written as

$$\frac{1-\lambda}{2} + \frac{1}{2|\mathcal{R}|}\#\left\{n\,\middle|\,\begin{matrix} n\in\{1,\ldots,L\} \\ \wedge|m-n|\in\mathcal{R}\wedge a_n = 1\end{matrix}\right\}$$

$$= \frac{1-\lambda}{2} + \frac{1}{2|\mathcal{R}|}\sum_{n=1}^{L}\delta_{a_n,1}\sum_{l=1}^{t}\left(\delta_{n,m+_L l} + \delta_{n,m-_L l}\right). \tag{C2}$$

Here, $\#\{n \mid A(n)\}$ denotes the number of n satisfying the condition $A(n)$, and $-_L$ denotes subtraction modulo L, namely, for integers (p,q) with $1 \leqslant p \leqslant L$ and $1 \leqslant q \leqslant L$,

$$p -_L q = \begin{cases} p-q & \text{if } p \geqslant q+1, \\ p-q+L & \text{if } p \leqslant q. \end{cases} \tag{C3}$$

Also, the coefficient of $|m\rangle_{\mathrm{B}\,\mathrm{B}}\langle n|$ for m,n such that $(a_m,a_n) = (1,1)$ in Eq. (C1) is written as

$$\begin{cases} \frac{\lambda}{2|\mathcal{R}|} & \text{if} \quad n = m \pm_L l (1 \leqslant l \leqslant t), \\ 0 & \text{otherwise.} \end{cases} \tag{C4}$$

Next, we cassify $\vec{a}$ in terms of the resulting eigenvalues of Eq. (C1). In so doing, note that the translation operation (as this is a unitary operator) defined by Eq. (47) does not change the eigenvalues of Eq. (C1). Hence, the eigenvalues of Eq. (C1) with $\vec{a}$ and $\vec{a}'$ are the same if there exists κ $(1 \leqslant \kappa \leqslant L)$ such that $a_k = a'_{k+_L\kappa}$ is satisfied for any $k \in \{1,\ldots,L\}$ and hence it suffices to consider $\vec{a} \in [\vec{a}] = \{\vec{a}' \mid \exists\kappa \in \{1,\ldots,L\}$ s.t. $a'_{k+_L\kappa} = a_k$ for $\forall k \in \{1,\ldots,L\}\}$ to derive the eigenvalues of Eq. (C1). In order to characterize $\vec{a}$, we introduce an $|\vec{a}|(= \nu+1)$-length vector $\vec{p} = (p_1 p_2 \ldots p_{\nu+1})$

that satisfies

$$a_{p_j} = 1 \quad \text{for } \forall j \in \{1,\ldots,\nu+1\}, \tag{C5}$$

$$1 \leqslant p_1 < p_2 < \cdots < p_{\nu+1} \leqslant L. \tag{C6}$$

By using $\vec{p}$, we can convert the problem of deriving the largest eigenvalue of Eq. (C1) to the maximization problem of the largest eigenvalue of the following matrix.

$$\frac{1-\lambda}{2} I_{\nu+1} + \frac{1}{2|\mathcal{R}|} B_{\nu+1}(\lambda). \tag{C7}$$

Here, $I_{\nu+1}$ denotes the $(\nu+1)\times(\nu+1)$ identity matrix and $B_{\nu+1}(\lambda)$ denotes the $(\nu+1)\times(\nu+1)$ matrix whose diagonal element $(B_{\nu+1}(\lambda))_{m,m}$ is given by

$$\begin{aligned}(B_{\nu+1}(\lambda))_{m,m} &= \#\{k \mid k \in \{1,\ldots,\nu+1\} \wedge |p_m - p_k| \in \mathcal{R}\} \\ &= \sum_{k=1}^{\nu+1}\sum_{l=1}^{t}\left(\delta_{p_k,p_m+_L l} + \delta_{p_k,p_m-_L l}\right) \\ &= \sum_{l=1}^{t}\sum_{k=1}^{\nu+1}\left(\delta_{p_k,p_m+_L l} + \delta_{p_k,p_m-_L l}\right) \\ &= \sum_{l=1}^{t}\sum_{k=1:k\neq m}^{\nu+1}\left(\delta_{p_k,p_m+_L l} + \delta_{p_k,p_m-_L l}\right),\end{aligned} \tag{C8}$$

and its off-diagonal element $(B_{\nu+1}(\lambda))_{m,n}$ $(m \neq n)$ is given by

$$(B_{\nu+1}(\lambda))_{m,n} = \begin{cases}\lambda & \text{if } p_n = p_m \pm_L l (1 \leqslant l \leqslant t), \\ 0 & \text{otherwise.}\end{cases} \tag{C9}$$

Among $[\vec{a}]$, we need to find $\vec{a}$ that achieves the largest eigenvalue of Eq. (C7). For this, we use the following fact.

Fact 1. For any real matrix with non-negative off-diagonal elements, the largest eigenvalue is maximized when all the matrix elements are maximized.

Proof. We consider two $n\times n$ real matrices $A = (A_{i,j})_{i,j}$ and $\tilde{A} = (\tilde{A}_{i,j})_{i,j}$ such that $A_{i,j} \geqslant \tilde{A}_{i,j}$ holds for any $i,j \in \{1,\ldots n\}$ and $A_{i,j}, \tilde{A}_{i,j} \geqslant 0$ holds if $i \neq j$. Suppose that $|\psi\rangle = (x_1 x_2 \ldots x_n)^{\mathrm{T}}$ and $|\tilde{\psi}\rangle = (\tilde{x}_1\tilde{x}_2\ldots\tilde{x}_n)^{\mathrm{T}}$ are normalized eigenvectors of A and $\tilde{A}$ that give the largest eigenvalue of A and $\tilde{A}$, respectively. Since both A and $\tilde{A}$ are real and all the off-diagonal elements of A and $\tilde{A}$ are non-negative, we can choose $|\psi\rangle$ and $|\tilde{\psi}\rangle$ such that all the elements of $|\psi\rangle$ and $|\tilde{\psi}\rangle$ are real and non-negative. By using $|\psi\rangle$ and $|\tilde{\psi}\rangle$, the largest eigenvalue of A and $\tilde{A}$ are, respectively, given by

$$\langle A\rangle_{\max} := \langle\psi|\, A\,|\psi\rangle = \sum_{i,j} A_{i,j} x_i x_j, \tag{C10}$$

$$\langle \tilde{A}\rangle_{\max} := \langle\tilde{\psi}|\, \tilde{A}\,|\tilde{\psi}\rangle = \sum_{i,j} \tilde{A}_{i,j} \tilde{x}_i \tilde{x}_j. \tag{C11}$$

Since $A_{i,j} \geqslant \tilde{A}_{i,j}$ holds for any $i,j \in \{1,\ldots n\}$ and $|\psi\rangle$ gives the largest eigenvalue of A, we have

$$\begin{aligned}\langle\tilde{A}\rangle_{\max} &= \sum_{i,j}\tilde{A}_{i,j}\tilde{x}_i\tilde{x}_j \leqslant \sum_{i,j} A_{i,j}\tilde{x}_i\tilde{x}_j = \langle\tilde{\psi}|\, A\,|\tilde{\psi}\rangle \\ &\leqslant \langle\psi|\, A\,|\psi\rangle = \langle A\rangle_{\max},\end{aligned} \tag{C12}$$

which ends the proof.

By using Fact 1, the largest eigenvalue of Eq. (C7) is obtained when $p_j + 1 = p_{j+1}$ for all $j \in \{1,\ldots,\nu\}$, namely, $\vec{a} = (0\ldots0\overbrace{1\ldots1}^{\nu+1}0\ldots0)$ in Eq. (C1). For example, if $\nu = 2$ and $|\mathcal{R}| = 2$, Eq. (C7) with $\vec{a} = (0\ldots01110\ldots0)$ is rewritten as

$$\begin{bmatrix} \frac{1-\lambda}{2}+\frac{1}{4} & \frac{\lambda}{4} & 0 \\ \frac{\lambda}{4} & \frac{1-\lambda}{2}+\frac{1}{2} & \frac{\lambda}{4} \\ 0 & \frac{\lambda}{4} & \frac{1-\lambda}{2}+\frac{1}{4} \end{bmatrix}, \tag{C13}$$

and this results in the largest eigenvalue of Eq. (C1), which corresponds to $\Omega_+^{(2)}(\lambda)$ for $\mathcal{R} = \{1, L\}$. Moreover, if $\nu \leqslant t(= |\mathcal{R}|/2)$, Eq. (C7) with $\vec{a} = (0\ldots0\overbrace{1\ldots1}^{\nu+1}0\ldots0)$ is rewritten as

$$\frac{1-\lambda}{2} I_{\nu+1} + \frac{1}{2|\mathcal{R}|}\mathcal{B}_{\nu+1}(\lambda), \tag{C14}$$

where $\mathcal{B}_{\nu+1}(\lambda)$ denotes the $(\nu+1)\times(\nu+1)$ matrix whose elements are given by

$$(\mathcal{B}_{\nu+1}(\lambda))_{m,n} = \begin{cases}\nu & \text{if } m = n, \\ \lambda & \text{otherwise.}\end{cases} \tag{C15}$$

Equation (C14) has only two eigenvalues: $(1-\lambda)/2 + (\nu-\lambda)/(2|\mathcal{R}|)$ and $(1-\lambda)/2 + \nu(1+\lambda)/(2|\mathcal{R}|)$. Since $\lambda \geqslant 0$, we have

$$\Omega_+^{(\nu)}(\lambda) = \frac{1-\lambda}{2} + \nu\frac{1+\lambda}{2|\mathcal{R}|}, \tag{C16}$$

which concludes Eq. (44).

APPENDIX D: PROOF OF THEOREM 1

Here, we prove Theorem 1 in the main text. In order to maximize $e_-^{(\mathrm{ph},1)}$, we derive an upper bound on $\Omega_-^{(1)}(\lambda)$. For this, we consider the largest eigenvalue of $\hat{\Pi}_{\vec{a}}^{(\mathrm{ph})} - \lambda\hat{\Pi}^{(\mathrm{b})}$ for $|\vec{a}| = \nu - 1 = 0$, namely, $\vec{a} = (00\ldots0) =: \vec{0}$. Since $\hat{\Pi}_{\vec{0}}^{(\mathrm{ph})} = 0$, $\Omega_-^{(1)}(\lambda)$ is given by the largest eigenvalue of $-\lambda\hat{\Pi}^{(\mathrm{b})}$, which is nonpositive. Hence, $e_-^{(\mathrm{ph},1)}$ is upper bounded by $e_-^{(\mathrm{ph},1)} = \min_{\lambda\geqslant0}\{\lambda e^{(\mathrm{b},1)} + \Omega_-^{(1)}(\lambda)\} \leqslant \min_{\lambda\geqslant0}\{\lambda e^{(\mathrm{b},1)}\} = 0$. For $e_+^{(\mathrm{b},1)}$, from Eq. (45), we obtain

$$e_+^{(\mathrm{ph},1)} \leqslant \frac{1}{2} + \frac{1}{2|\mathcal{R}|}, \tag{D1}$$

for $e^{(\mathrm{b},1)} \geqslant \frac{|\mathcal{R}|-1}{2|\mathcal{R}|}$. Also, for $0 \leqslant e^{(\mathrm{b},1)} \leqslant \frac{|\mathcal{R}|-1}{2|\mathcal{R}|}$, $e^{(\mathrm{ph},1)}$ is upper bounded by

$$e^{(\mathrm{ph},1)} \leqslant \frac{2|\mathcal{R}|}{|\mathcal{R}|-1} e^{(\mathrm{b},1)}\left(\frac{1}{2} + \frac{1}{2|\mathcal{R}|}\right) = \frac{|\mathcal{R}|+1}{|\mathcal{R}|-1} e^{(\mathrm{b},1)}, \tag{D2}$$

by choosing p in Eq. (52) as $p = \frac{2|\mathcal{R}|}{|\mathcal{R}|-1} e^{(\mathrm{b},1)}$. Therefore, by combining Eqs. (D1) and (D2), we conclude Eq. (50).

[1] C. Bennett and G. Brassard, Advances in *Proceedings of the IEEE International Conference on Computers*, Systems and Signal Processing (IEEE, Bangalore, India, 1984), p. 175.
[2] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, Phys. Rev. Lett. **67**, 661 (1991).
[3] C. H. Bennett, Quantum Cryptography Using any two Nonorthogonal States, Phys. Rev. Lett. **68**, 3121 (1992).
[4] D. Bruß, Optimal Eavesdropping in Quantum Cryptography with Six States, Phys. Rev. Lett. **81**, 3018 (1998).
[5] K. Inoue, E. Waks, and Y. Yamamoto, Differential Phase Shift Quantum Key Distribution, Phys. Rev. Lett. **89**, 037902 (2002).
[6] V. Scarani, A. Acin, G. Ribordy, and N. Gisin, Quantum Cryptography Protocols Robust Against Photon Number Splitting Attacks for Weak Laser Pulse Implementations, Phys. Rev. Lett. **92**, 057901 (2004).
[7] N. Gisin *et al.*, Towards practical and fast quantum cryptography, arXiv:quant-ph/0411022.
[8] M. Sasaki *et al.*, Field test of quantum key distribution in the Tokyo QKD network, Opt. Express **19**, 10387 (2011).
[9] K. Tamaki, M. Koashi, and G. Kato, Unconditional security of coherent-state-based differential phase shift quantum key distribution protocol with block-wise phase randomization, arXiv:1208.1995.
[10] Note that the security of the coherent-one-way protocol, which is similar to the DPS protocol, was analyzed in T. Moroder, M. Curty, C. C. W. Lim, L. P. Thinh, H. Zbinden, and N. Gisin, Security of Distributed-Phase-Reference Quantum Key Distribution, Phys. Rev. Lett. **109**, 260501 (2012).
[11] T. Sasaki, Y. Yamamoto, and M. Koashi, Practical quantum key distribution protocol without monitoring signal disturbance, Nature (London) **509**, 475 (2014).
[12] H. Takesue, T. Sasaki, K. Tamaki, and M. Koashi, Experimental quantum key distribution without monitoring signal disturbance, Nat. Photonics **9**, 827 (2015).
[13] A. Mizutani, N. Imoto, and K. Tamaki, Robustness of the round-robin differential-phase-shift quantum-key-distribution protocol against source flaws, Phys. Rev. A **92**, 060303(R) (2015).
[14] Z. Zhang, X. Yuan, Z. Cao, and X. Ma, Practical round-robin differential-phase-shift quantum key distribution, New J. Phys. **19**, 033013 (2017).
[15] H. F. Chau, C. Wong, Q. Wang, and T. Huang, Qudit-based measurement-device-independent quantum key distribution using linear optics, arXiv:1608.08329.
[16] T. Sasaki, K. Tamaki, and M. Koashi, Quantum key distribution protocols with slow basis choice, arXiv:1604.04460.
[17] Y.-H. Li *et al.*, Experimental round-robin differential phase-shift quantum key distribution, Phys. Rev. A **93**, 030302(R) (2016).
[18] S. Wang *et al.*, Experimental demonstration of a quantum key distribution without signal disturbance monitoring, Nat. Photonics **9**, 832 (2015).
[19] J.-Y. Guan *et al.*, Experimental Passive Round-Robin Differential Phase-Shift Quantum Key Distribution, Phys. Rev. Lett. **114**, 180502 (2015).
[20] P. W. Shor and J. Preskill, Simple Proof of Security of the BB84 Quantum Key Distribution Protocol, Phys. Rev. Lett. **85**, 441 (2000).
[21] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, Security of quantum key distribution with imperfect devices, Quantum Inf. Comput. **4**, 325 (2004).