

This is the accepted manuscript made available via CHORUS. The article has been published as:

# Finite-key analysis for quantum key distribution with weak coherent pulses based on Bernoulli sampling

Shun Kawakami, Toshihiko Sasaki, and Masato Koashi

Phys. Rev. A **96**, 012305 — Published 5 July 2017

DOI: [10.1103/PhysRevA.96.012305](https://doi.org/10.1103/PhysRevA.96.012305)

# Finite-key analysis for quantum key distribution with weak coherent pulses based on Bernoulli sampling

Shun Kawakami<sup>1,2</sup>, Toshihiko Sasaki<sup>1</sup> and Masato Koashi<sup>1,2</sup>

<sup>1</sup>*Photon Science Center, Graduate School of Engineering,  
The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8656, Japan*

<sup>2</sup>*Department of Applied Physics, Graduate School of Engineering,  
The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

An essential step in quantum key distribution is the estimation of parameters related to the leaked amount of information, which is usually done by sampling of the communication data. When the data size is finite, the final key rate depends on how the estimation process handles statistical fluctuations. Many of the present security analyses are based on the method with simple random sampling, where hypergeometric distribution or its known bounds are used for the estimation. Here we propose a concise method based on Bernoulli sampling, which is related to binomial distribution. Our method is suitable for the BB84 protocol with weak coherent pulses, reducing the number of estimated parameters to achieve a higher key generation rate compared to the method with simple random sampling. We also applied the method to prove the security of the differential-quadrature-phase-shift (DQPS) protocol in the finite-key regime. The result indicates that, the advantage of the DQPS protocol over the phase-encoding BB84 protocol in terms of the key rate, which was previously confirmed in the asymptotic regime, persists in the finite-key regime.

PACS numbers: 03.67.Dd, 03.67.Hk.

## I. INTRODUCTION

Quantum key distribution (QKD) allows two distant parties to share a secret key and realizes a communication with information-theoretic security by combining it with one-time-pad encryption. Since the BB84 protocol was proposed by Bennett and Brassard [1], a large number of researches on QKD have been conducted from both aspects of theory and implementations. The security of QKD is based only on principles of quantum physics where eavesdropped information is bounded from the observed parameters in a protocol. In practice, the estimation of this bound should take into account statistical fluctuations due to the finite size of communication data, which requires so-called finite-key analysis. Following the security definition with compositability [2, 3], the finite-key analyses for various QKD protocols including the BB84 protocol were conducted assuming the adversary's general attacks [4–10].

For a finite-key analysis, a simple method with a smaller number of estimation processes is preferred because it leads not only to a more concise security proof but also to a higher key-generation rate especially when the number of communication rounds is limited and statistical fluctuations are large. A number of recent finite-key analyses are based on the method with simple random sampling, which is used to model  $n_1$  draws, without replacement, from a finite population of size  $n_2$  that contains  $k_2$  errors. The probability that the number of errors in the sample is  $k_1$  obeys hypergeometric distribution

$$\text{HG}(k_1; n_1, k_2, n_2) := \frac{\binom{k_2}{k_1} \binom{n_2 - k_2}{n_1 - k_1}}{\binom{n_2}{n_1}}. \quad (1)$$

In several finite-key analyses [8, 9] based on simple random sampling, efforts were made to find bounds on hypergeometric distribution which are related to binomial distribution in order to simplify numerical calculation.

In order to mitigate the inefficiency arising from basis mismatch between the sender and the receiver, the BB84 protocol is often implemented with biased basis choice [11], in which the minor basis is used solely for monitoring leaked information in the major basis. In this case, the whole data from the rounds in the monitoring basis is regarded as a sample, with each round selected with a constant probability dictated in the protocol as that of the basis choice. This suggests that the data from the monitoring basis is related to Bernoulli sampling, in which each element of the population of size  $n_2$  is sampled with fixed probability  $p_1$ . The number of samples  $n_1$  obeys binomial distribution

$$\text{BI}(n_1; n_2, p_1) := \binom{n_2}{n_1} p_1^{n_1} (1 - p_1)^{n_2 - n_1}. \quad (2)$$

If the BB84 protocol with biased basis choice essentially includes the property of the binomial distribution, analysis based on the conventional simple random sampling may introduce unnecessary complexity and possibly leads to a lower key rate.

In this paper, we work on the finite-key analysis by focusing on the Bernoulli sampling instead of simple random sampling. We propose the method based on binomial distribution which is parametrized by the basis choice probabilities in the protocol. Differently from the previous works which deal with binomial distribution to derive bounds on hypergeometric distribution, our work is based on binomial distributions inherent in the protocol. Our method is especially suited for the BB84 protocol with weak coherent pulses (WCP), providing a simpler analysis with less estimation processes as well as achieving a higher key rate compared to the analysis with simple random sampling. We further apply this method to the differential quadrature phase shift (DQPS) protocol [12], whose security was proved recently in the asymptotic regime [13]. As a result, we show that the advantage of the DQPS protocol over the phase-encoded BB84 protocol with WCP still

remains in the finite key regime.

This paper is organized as follows. In Sec. II, we describe details of the BB84 protocol which is considered in this work, along with a summary of notations used in this paper. In Sec. III, we propose a method of finite-key analysis based on Bernoulli sampling, and applies it to the ideal BB84 protocol where Alice and Bob can manipulate perfect single photon states. The proposed method is then applied to the BB84 protocol with WCP as well as the DQPS protocol in Sec. IV A. Finally, we give discussion and conclusion in Sec. V.

## II. BB84 PROTOCOL

In most part of this paper, we discuss the finite key analysis of the BB84 protocol [14], which is given in the following. The sender Alice and the receiver Bob independently chooses two bases ( $Z$  basis and  $X$  basis) with a biased probability. The final key is generated only from  $Z$ -basis data, while  $X$ -basis data is used for leak monitoring to determine the amount for privacy amplification. The number of total rounds  $n_{\text{rep}}$  are predetermined and there is no threshold for data size after sifting process, which means that the size of sifted key  $n_Z$  and that of monitoring bits  $n_X$  are not determined until quantum communication is over.

The protocol proceeds as follows with predetermined parameters  $\tilde{p}_Z$ ,  $\tilde{p}_X = 1 - \tilde{p}_Z$  and  $n_{\text{rep}}$ . In its description,  $|k|$  represents the length of a bit sequence  $k$ .

- (1) Alice chooses  $Z$  basis or  $X$  basis with probability  $\tilde{p}_Z$  and  $\tilde{p}_X$ , respectively. She chooses a uniformly random bit  $\{0, 1\}$ .
- (2) Alice prepares one of states  $\{\hat{\rho}_{Z,0}, \hat{\rho}_{Z,1}, \hat{\rho}_{X,0}, \hat{\rho}_{X,1}\}$  based on the selected basis and bit. She sends the prepared state to Bob over the quantum channel.
- (3) Bob chooses  $Z$  basis or  $X$  basis with probability  $\tilde{p}_Z$  and  $\tilde{p}_X$ , respectively. He measures a received state in chosen basis and obtains the outcome  $\{0, 1, \text{no-detection}\}$ .
- (4) They repeat the sequence (1) to (3), which we call a round, by  $n_{\text{rep}}$  times.
- (5) Bob publicly announces whether each round has resulted in a detection or not. Let  $n_{\text{det}}$  be the number of rounds with detection.
- (6) Alice and Bob disclose all of their basis choices. Among the  $n_{\text{det}}$  detected rounds, the rounds where both Alice and Bob chose the  $Z$  basis are called “ $Z$ -labeled” rounds, and the rounds where they chose the  $X$  basis are called “ $X$ -labeled” rounds. They define sifted keys  $\kappa_{A,Z}$  and  $\kappa_{B,Z}$  by concatenating the bits for the  $Z$ -labeled rounds, and similarly define  $\kappa_{A,X}$  and  $\kappa_{B,X}$  for the  $X$ -labeled rounds. Let their sizes be  $n_Z := |\kappa_{A,Z}| = |\kappa_{B,Z}|$  and  $n_X := |\kappa_{A,X}| = |\kappa_{B,X}|$ .
- (7) They disclose and compare  $\kappa_{A,X}$  and  $\kappa_{B,X}$  to determine the number of bit errors  $k_X$  included in them.
- (8) Through public discussion, Bob corrects his keys  $\kappa_{B,Z}$  to make it coincide with Alice’s key  $\kappa_{A,Z}$  and obtains  $\kappa_{B,Z}^{\text{cor}}$  ( $|\kappa_{B,Z}^{\text{cor}}| = n_Z$ ).
- (9) Alice and Bob conduct privacy amplification by shortening  $\kappa_{A,Z}$  and  $\kappa_{B,Z}^{\text{cor}}$  to obtain final keys  $\kappa_{A,Z}^{\text{fin}}$  and  $\kappa_{B,Z}^{\text{fin}}$  of size  $n_{\text{fin}}$ .

In the subsequent sections, we discuss how we can determine the final key length  $n_{\text{fin}}$  as a function of the random variables  $n_Z$ ,  $n_X$ , and  $k_X$  obtained in the protocol to satisfy a given security criteria. For convenience, we define several variables and parameters as  $n_{\text{tot}} := n_X + n_Z$  and

$$\begin{aligned} p_Z &:= \tilde{p}_Z^2 / (\tilde{p}_Z^2 + \tilde{p}_X^2), \\ p_X &:= \tilde{p}_X^2 / (\tilde{p}_Z^2 + \tilde{p}_X^2). \end{aligned} \quad (3)$$

Throughout this paper, we adopt an abuse of notation to use the same symbol for a random variable  $\tilde{n}$  and its value  $n$ , whenever the distinction is obvious. For example, we denote  $\Pr(n > 3)$  instead of  $\Pr(\tilde{n} > 3)$ . We denote by  $\Pr(n)$  the probability mass function  $\Pr(\tilde{n} = n)$ . Similarly, we use  $\Pr(n | m)$  instead of  $\Pr(\tilde{n} = n | \tilde{m} = m)$ . We define  $|0_Z\rangle$  and  $|1_Z\rangle$  as basis vectors of  $Z$  basis on a qubit system, and  $|0_X\rangle := (|0_Z\rangle + |1_Z\rangle)/\sqrt{2}$  and  $|1_X\rangle := (|0_Z\rangle - |1_Z\rangle)/\sqrt{2}$  as those of  $X$  basis. When the same notations are used for an optical signal, it should be understood that they refer to the states in the subspace of a single photon contained in two modes, such as polarizations. The four Bell states are represented as  $|\Phi^\pm\rangle$  and  $|\Psi^\pm\rangle$  where

$$|\Phi^\pm\rangle := \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \quad (4)$$

$$|\Psi^\pm\rangle := \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle). \quad (5)$$

We define a function  $h(x)$  for  $x \geq 0$  as

$$h(x) = \begin{cases} -x \log_2 x - (1-x) \log_2 (1-x) & (0 \leq x \leq 1/2) \\ 1 & (x > 1/2). \end{cases} \quad (6)$$

## III. ANALYSIS FOR THE IDEAL BB84 PROTOCOL

Here we consider finite-key analysis for the ideal qubit-based BB84 protocol, in which Alice sends a single photon in the states  $\{\hat{\rho}_{W,a} = |a_W\rangle\langle a_W|_S\}$  ( $W \in \{Z, X\}, a \in \{0, 1\}$ ) in Step (2) and Bob conducts ideal measurement with unit efficiency described by POVM (positive operator-valued measure)  $\{|0_W\rangle\langle 0_W|_S, |1_W\rangle\langle 1_W|_S, \mathbb{1}_S - |0_W\rangle\langle 0_W|_S - |1_W\rangle\langle 1_W|_S\}$  corresponding to the outcome  $\{0, 1, \text{no-detection}\}$ .

### A. Security criteria and formalism for key length

In this work, we follow the security definition based on universally composable security [2, 3]. The protocol is called  $\epsilon_{\text{sec}}$  secure if it is both  $\epsilon_c$ -correct and  $\epsilon_s$ -secret where  $\epsilon_{\text{sec}} = \epsilon_c + \epsilon_s$ . We call the protocol is  $\epsilon_c$ -correct if  $\Pr(\kappa_{A,Z}^{\text{fin}} \neq \kappa_{B,Z}^{\text{fin}}) \leq \epsilon_c$  holds. We call the protocol is  $\epsilon_s$ -secret if

$$\frac{1}{2} \|\hat{\rho}_{AE}^{\text{fin}} - \hat{\rho}_{AE}^{\text{ideal}}\| \leq \epsilon_s, \quad (7)$$

where  $\hat{\rho}_{AE}^{\text{fin}} := \sum_{\kappa_{A,Z}^{\text{fin}}} p(\kappa_{A,Z}^{\text{fin}}) |\kappa_{A,Z}^{\text{fin}}\rangle\langle \kappa_{A,Z}^{\text{fin}}| \otimes \hat{\rho}_E(\kappa_{A,Z}^{\text{fin}})$  is a classical-quantum state between Alice’s key and Eve’s system after

finishing the protocol and  $\hat{\rho}_{AE}^{\text{ideal}}$  is an ideal separable state in which Alice's key is uniformly distributed over  $2^{|k_{AZ}^{\text{fin}}|}$  values and decoupled from Eve's system.

In the ideal BB84 protocol, Alice's procedure of selecting a random bit and a basis, and preparing the corresponding signal can be replaced [15] by preparation of  $|\Phi^+\rangle_{AS}$  followed by the measurement on the system  $A$  on  $\{|0_Z\rangle_A, |1_Z\rangle_A\}$  ( $Z$ -basis) or on  $\{|0_X\rangle_A, |1_X\rangle_A\}$  ( $X$ -basis). Bob's measurement is also replaced by a filtering operation to make sure a single photon is received and transfer its state to a qubit  $B$ , followed by the orthogonal measurement of  $B$  on the chosen basis to determine the outcome 0 or 1. When the filtering fails, the outcome is "no-detection". According to Ref. [16, 17], a phase error occurs when Alice and Bob conduct virtual Bell-basis measurement on a  $Z$ -labeled round after Eve's intervention to obtain the outcome for  $|\Phi^-\rangle\langle\Phi^-|_{AB}$  or  $|\Psi^-\rangle\langle\Psi^-|_{AB}$ . Since we have the relation

$$|\Phi^-\rangle\langle\Phi^-|_{AB} + |\Psi^-\rangle\langle\Psi^-|_{AB} = |01_X\rangle\langle 01_X|_{AB} + |10_X\rangle\langle 10_X|_{AB}, \quad (8)$$

phase error is equivalently defined as a bit error which occurs when Alice and Bob conduct virtual  $X$ -basis measurement after Eve's intervention on a  $Z$ -labeled round. An important property which will be used in the next subsection is that the measurement for a phase error on a  $Z$ -labeled round and the measurement for a bit error on an  $X$ -labeled round are identical, and hence they only differs in the labeling.

Let  $k_{\text{ph}}$  be a random variable which represents the number of phase errors on  $n_Z$   $Z$ -labeled rounds. Once we have a good upper bound on  $k_{\text{ph}}$ , a secure key length can be calculated as follows. Suppose that we have a function  $f(k_X, n_X, n_{\text{tot}})$  which satisfies

$$\Pr(k_{\text{ph}} > f(k_X, n_X, n_{\text{tot}})) \leq \epsilon_{\text{PE}} \quad (9)$$

regardless of Eve's attack strategy. By setting [8]

$$\epsilon_s = \sqrt{2} \sqrt{\epsilon_{\text{PE}} + \epsilon_{\text{PA}}}, \quad (10)$$

the protocol is  $\epsilon_c$ -correct and  $\epsilon_s$ -secret if the final key length  $n_{\text{fin}}$  satisfies

$$n_{\text{fin}} \leq n_Z - \left\lceil n_Z h \left( \frac{f(k_X, n_X, n_{\text{tot}})}{n_Z} \right) + \log_2 \frac{1}{\epsilon_{\text{PA}}} \right\rceil - \lambda_{\text{EC}}(\epsilon_c), \quad (11)$$

where  $\lceil \cdot \rceil$  represents the ceiling function and  $\lambda_{\text{EC}}(\epsilon_c)$  is the cost of error correction to achieve  $\epsilon_c$ -correctness. For simplicity, we will replace the right-hand side by a slightly pessimistic bound as

$$n_{\text{fin}} \leq n_Z (1 - h \left( \frac{f(k_X, n_X, n_{\text{tot}})}{n_Z} \right)) - \log_2 \frac{2}{\epsilon_{\text{PA}}} - \lambda_{\text{EC}}(\epsilon_c). \quad (12)$$

## B. Bounds on phase errors

In this subsection, we discuss the specific methods to obtain  $f(k_X, n_X, n_{\text{tot}})$  in Eq. (9) including a method based on the Bernoulli sampling, and a more conventional method based on

the simple random sampling. We also introduce a third, rather convoluted method, which will help to elucidate the difference between the former two methods.

Before discussing each of the methods, we first derive general statistical properties. Since the  $Z$ -labeled phase error and the  $X$ -labeled bit error are obtained by identical measurements, the procedure to obtain those errors is equivalent to the following steps after discarding the rounds with no-detection (i.e., with Bob failing to receive a qubit): (a) Alice and Bob further discard each of the remaining rounds with probability  $1 - \tilde{p}_Z^2 - \tilde{p}_X^2$ . (b) They make  $X$ -basis measurements on the remaining  $n_{\text{tot}}$  rounds and obtain  $k_{\text{tot}}$  errors. (c) Finally, they label each of the  $n_{\text{tot}}$  rounds as  $Z$  or  $X$  with probability  $p_Z$  and  $p_X$  (see Eq. (3)), respectively, and obtain  $k_{\text{ph}}$  phase errors in  $Z$ -labeled rounds and  $k_X = k_{\text{tot}} - k_{\text{ph}}$  bit errors in  $X$ -labeled rounds. In this procedure, since  $k_X$  errors are sampled from  $k_{\text{tot}}$  errors with a fixed probability  $p_X$ , it follows a binomial distribution if  $k_{\text{tot}}$  and  $n_{\text{tot}}$  are fixed:

$$\Pr(k_X | k_{\text{tot}}, n_{\text{tot}}) = \text{BI}(k_X; k_{\text{tot}}, p_X). \quad (13)$$

On the other hand, the step (c) of the above procedure is equivalently denoted as follows: Alice and Bob draw a number  $n_X$  based on the binomial distribution  $\text{BI}(n_X; n_{\text{tot}}, p_X)$ , and then select  $n_X$  random rounds among the  $n_{\text{tot}}$  rounds to label as  $X$ , thereby determining  $k_X$ . This implies that the number  $k_X$  obeys hypergeometric distribution if  $n_X, k_{\text{tot}}$  and  $n_{\text{tot}}$  are fixed:

$$\Pr(k_X | n_X, k_{\text{tot}}, n_{\text{tot}}) = \text{HG}(k_X; n_X, k_{\text{tot}}, n_{\text{tot}}). \quad (14)$$

In order to use the properties derived above, it is convenient to reformulate Eq. (9) as follows. From Eq. (9), we have

$$\sum_{k_{\text{tot}}, n_{\text{tot}}} \Pr(k_{\text{ph}} > f(k_X, n_X, n_{\text{tot}}) | k_{\text{tot}}, n_{\text{tot}}) \Pr(k_{\text{tot}}, n_{\text{tot}}) \leq \epsilon_{\text{PE}}. \quad (15)$$

Since  $\Pr(k_{\text{tot}}, n_{\text{tot}})$  can be under control of Eve, we seek for  $f(k_X, n_X, n_{\text{tot}})$  satisfying

$$\Pr(k_{\text{ph}} > f(k_X, n_X, n_{\text{tot}}) | k_{\text{tot}}, n_{\text{tot}}) \leq \epsilon_{\text{PE}} \quad (16)$$

for any  $k_{\text{tot}}$  and  $n_{\text{tot}}$ , which is a sufficient condition for Eq. (9). For later convenience, we equivalently describe Eq. (16) as

$$\sum_{k_X, n_X; k_X < k_{\text{tot}} - f(k_X, n_X, n_{\text{tot}})} \Pr(k_X, n_X | k_{\text{tot}}, n_{\text{tot}}) \leq \epsilon_{\text{PE}}. \quad (17)$$

The first method to determine  $f(k_X, n_X, n_{\text{tot}})$ , whose utility we will emphasize throughout this paper, is based on Bernoulli sampling using the property of binomial distribution Eq. (13). This method adopts  $f(k_X, n_X, n_{\text{tot}}) = f_{\text{BI}}(k_X)$  where

$$f_{\text{BI}}(k_X) := \min \left\{ k_{\text{tot}} \mid C_{\text{BI}}(k_X; k_{\text{tot}}, p_X) \leq \epsilon_{\text{PE}} \right\} - k_X - 1 \quad (18)$$

$$C_{\text{BI}}(k_X; k_{\text{tot}}, p_X) := \sum_{k'_X \leq k_X} \text{BI}(k'_X; k_{\text{tot}}, p_X). \quad (19)$$

The proof that  $f_{\text{BI}}(k_X)$  satisfies Eq. (16) is as follows. Let  $\bar{k}_X(k_{\text{tot}}) := \max\{k_X | k_{\text{tot}} > f_{\text{BI}}(k_X) + k_X\}$ . Then we have

$$\sum_{k_X; k_{\text{tot}} > f_{\text{BI}}(k_X) + k_X} \text{BI}(k_X; k_{\text{tot}}, p_X) \leq C_{\text{BI}}(\bar{k}_X(k_{\text{tot}}); k_{\text{tot}}, p_X). \quad (20)$$

Since  $C_{\text{BI}}(k_X; k_{\text{tot}}, p_X)$  is a decreasing function of  $k_{\text{tot}}$ , from Eq. (18) we have  $C_{\text{BI}}(k_X; k_{\text{tot}}, p_X) \leq \epsilon_{\text{PE}}$  for any pair  $(k_X, k_{\text{tot}})$  satisfying  $k_{\text{tot}} \geq f_{\text{BI}}(k_X) + k_X + 1$ . Since  $k_{\text{tot}} \geq f_{\text{BI}}(\bar{k}_X(k_{\text{tot}})) + \bar{k}_X(k_{\text{tot}}) + 1$  holds by definition of  $\bar{k}_X(k_{\text{tot}})$ , we have  $C_{\text{BI}}(\bar{k}_X(k_{\text{tot}}); k_{\text{tot}}, p_X) \leq \epsilon_{\text{PE}}$ . By connecting this to Eq. (20), we have

$$\sum_{k_X; k_X < k_{\text{tot}} - f_{\text{BI}}(k_X)} \text{BI}(k_X; k_{\text{tot}}, p_X) \leq \epsilon_{\text{PE}}, \quad (21)$$

for any  $k_{\text{tot}}$ . From Eqs. (13) and (21), we have

$$\begin{aligned} & \sum_{k_X, n_X; k_X < k_{\text{tot}} - f_{\text{BI}}(k_X)} \Pr(k_X, n_X \mid k_{\text{tot}}, n_{\text{tot}}) \\ &= \sum_{k_X; k_X < k_{\text{tot}} - f_{\text{BI}}(k_X)} \Pr(k_X \mid k_{\text{tot}}, n_{\text{tot}}) \\ &\leq \epsilon_{\text{PE}}, \end{aligned} \quad (22)$$

which is identical to Eq. (17) with  $f(k_X, n_X, n_{\text{tot}}) = f_{\text{BI}}(k_X)$ . Therefore, we have

$$\Pr(k_{\text{ph}} > f_{\text{BI}}(k_X) \mid k_{\text{tot}}, n_{\text{tot}}) \leq \epsilon_{\text{PE}}. \quad (23)$$

As a result of the Bernoulli-sampling method, the protocol is  $\epsilon_c$ -correct and  $\epsilon_s$ -secret if the final key length  $n_{\text{fin}}$  satisfies

$$n_{\text{fin}} \leq l^{(\text{BI})} := n_Z(1 - h\left(\frac{f_{\text{BI}}(k_X)}{n_Z}\right)) - \log_2 \frac{2}{\epsilon_{\text{PA}}} - \lambda_{\text{EC}}(\epsilon_c), \quad (24)$$

where  $\epsilon_s$  is given by Eq. (10).

The second method is based on simple random sampling, applying the property of the hypergeometric distribution Eq. (14), which is already seen in Ref. [4, 6–8], for example. This method adopts  $f(k_X, n_X, n_{\text{tot}}) = f_{\text{HG}}(k_X, n_X, n_{\text{tot}})$  where

$$\begin{aligned} f_{\text{HG}}(k_X, n_X, n_{\text{tot}}) &:= \min \left\{ k_{\text{tot}} \mid C_{\text{HG}}(k_X; n_X, k_{\text{tot}}, n_{\text{tot}}) \leq \epsilon_{\text{PE}} \right\} \\ &\quad - k_X - 1 \\ C_{\text{HG}}(k_X; n_X, k_{\text{tot}}, n_{\text{tot}}) &:= \sum_{k'_X \leq k_X} \text{HG}(k'_X; n_X, k_{\text{tot}}, n_{\text{tot}}). \end{aligned} \quad (25)$$

The proof that  $f_{\text{HG}}(k_X, n_X, n_{\text{tot}})$  satisfies Eq. (16) is similar to the proof for  $f_{\text{BI}}(k_X)$ . Recall that the proof for  $f_{\text{BI}}(k_X)$  did not use the explicit form of  $\text{BI}(k'_X, k_{\text{tot}}, p_X)$  but only used the decreasing property of  $C_{\text{BI}}(k_X; k_{\text{tot}}, p_X)$  as a function of  $k_{\text{tot}}$ . Since  $C_{\text{HG}}(k_X; n_X, k_{\text{tot}}, n_{\text{tot}})$  is also a decreasing function of  $k_{\text{tot}}$ , we have

$$\sum_{k_X; k_X < k_{\text{tot}} - f_{\text{HG}}(k_X, n_X, n_{\text{tot}})} \text{HG}(k_X; n_X, k_{\text{tot}}, n_{\text{tot}}) \leq \epsilon_{\text{PE}} \quad (26)$$

for any  $n_X, k_{\text{tot}}$  and  $n_{\text{tot}}$ , which is analogous to Eq. (21). From Eqs. (14) and (26), we have

$$\begin{aligned} & \sum_{k_X, n_X; k_X < k_{\text{tot}} - f_{\text{HG}}(k_X, n_X, n_{\text{tot}})} \Pr(k_X, n_X \mid k_{\text{tot}}, n_{\text{tot}}) \\ &= \sum_{k_X, n_X; k_X < k_{\text{tot}} - f_{\text{HG}}(k_X, n_X, n_{\text{tot}})} \Pr(k_X \mid n_X, k_{\text{tot}}, n_{\text{tot}}) \Pr(n_X \mid k_{\text{tot}}, n_{\text{tot}}) \\ &\leq \epsilon_{\text{PE}}, \end{aligned} \quad (27)$$

which is identical to Eq. (17) with  $f(k_X, n_X, n_{\text{tot}}) = f_{\text{HG}}(k_X, n_X, n_{\text{tot}})$ . Therefore, we have

$$\Pr(k_{\text{ph}} > f_{\text{HG}}(k_X, n_X, n_{\text{tot}}) \mid k_{\text{tot}}, n_{\text{tot}}) \leq \epsilon_{\text{PE}}. \quad (28)$$

As a result of the method with simple random sampling, the protocol is  $\epsilon_c$ -correct and  $\epsilon_s$ -secret if the secret key length  $n_{\text{fin}}$  satisfies

$$n_{\text{fin}} \leq l^{(\text{HG})} := n_Z(1 - h\left(\frac{f_{\text{HG}}(k_X, n_X, n_{\text{tot}})}{n_Z}\right)) - \log_2 \frac{2}{\epsilon_{\text{PA}}} - \lambda_{\text{EC}}(\epsilon_c), \quad (29)$$

where  $\epsilon_s$  is given by Eq. (10).

To understand the relation between the two methods with Bernoulli sampling and simple random sampling, we introduce another method which uses full knowledge of the distribution  $\Pr(k_X, n_X \mid k_{\text{tot}}, n_{\text{tot}})$  appearing in Eq. (17). The argument before Eq. (13) also implies that the number  $m_X := n_X - k_X$  of  $X$ -labeled rounds without bit error obeys binomial distribution  $\text{BI}(m_X; n_{\text{tot}} - k_{\text{tot}}, p_X)$ , and that  $m_X$  and  $k_X$  are independent conditioned on  $k_{\text{tot}}$  and  $n_{\text{tot}}$ . We thus obtain

$$\Pr(k_X, n_X \mid k_{\text{tot}}, n_{\text{tot}}) = \text{BI}(k_X; k_{\text{tot}}, p_X) \text{BI}(n_X - k_X; n_{\text{tot}} - k_{\text{tot}}, p_X). \quad (30)$$

The argument leading to Eq. (14) gives another expression for the distribution as

$$\Pr(k_X, n_X \mid k_{\text{tot}}, n_{\text{tot}}) = \text{HG}(k_X; n_X, k_{\text{tot}}, n_{\text{tot}}) \text{BI}(n_X; n_{\text{tot}}, p_X). \quad (31)$$

As a result, Eq. (17) is expressed in the following two equivalent ways:

$$\begin{aligned} & \sum_{k_X, m_X; k_X < k_{\text{tot}} - f(k_X, k_X + m_X, n_{\text{tot}})} \text{BI}(k_X; k_{\text{tot}}, p_X) \text{BI}(m_X; n_{\text{tot}} - k_{\text{tot}}, p_X) \\ &\leq \epsilon_{\text{PE}}. \end{aligned} \quad (32)$$

or

$$\sum_{k_X, n_X; k_X < k_{\text{tot}} - f(k_X, n_X, n_{\text{tot}})} \text{HG}(k_X; n_X, k_{\text{tot}}, n_{\text{tot}}) \text{BI}(n_X; n_{\text{tot}}, p_X) \leq \epsilon_{\text{PE}}. \quad (33)$$

Since  $f_{\text{BI}}(k_X)$  satisfies Eq. (21), Eq. (32) holds if  $f(k_X, k_X + m_X, n_{\text{tot}}) = f_{\text{BI}}(k_X)$ . Similarly, since  $f_{\text{HG}}(k_X, n_X, n_{\text{tot}})$  satisfies Eq. (26), Eq. (33) holds if  $f(k_X, n_X, n_{\text{tot}}) = f_{\text{HG}}(k_X, n_X, n_{\text{tot}})$ . On the other hand, the condition of Eqs. (32) and (33) do not imply Eq. (21) or Eq. (26). Therefore, there could be a better bound compared to  $f_{\text{BI}}(k_X)$  and  $f_{\text{HG}}(k_X, n_X, n_{\text{tot}})$  based on Eq. (32) or Eq. (33). In general, it is very complicated to determine the optimal function  $f(k_X, n_X, n_{\text{tot}})$  for the final key length  $n_{\text{fin}}$ , since it will depend on the explicit functional dependence of  $n_{\text{fin}}$  on  $f(k_X, n_X, n_{\text{tot}})$ .

The difference between the two equivalent conditions Eqs. (32) and (33) is the choice of two variables from three no-independent random variables  $k_X$ ,  $n_X$  and  $m_X$ . When  $(k_X, n_X)$  are chosen in Eq. (33), the distribution of  $k_X$ ,  $\text{HG}(k_X; n_X, k_{\text{tot}}, n_{\text{tot}})$  is dependent on the value of  $n_X$ . On the



other hand, Eq. (32) implies that two variables  $(k_X, m_X)$  are independent of each other. This suggests that the underlying statistics in the BB84 protocol with biased basis choice are understood in terms of independent binomial distributions.

Let us mention the difference from the other works [8, 9] which deal with relations between bounds on binomial distribution and ones on hypergeometric distribution since the former are easily treated with existing mathematical packages. Ref. [8] uses the property, which dates back to Hoeffding [18], that expectation of a convex function over hypergeometric distribution is no larger than that over binomial distribution. In [9], Ahrens map [19] was used to show that hypergeometric distribution is bounded by a permuted binomial distribution within a factor of  $\sqrt{2}$ . In contrast to these works, in our case the probability distribution Eq. (13) reflects the binomial distribution inherent in the BB84 protocol with biased basis choice.

### C. Numerical examples

Here we numerically compare the final key lengths derived from the three methods in the last subsection in the simplest cases. We calculate the key lengths for the case where no error is observed ( $k_X = 0$ ) and every signal is detected ( $n_{\text{tot}} = n_{\text{rep}}$ ). The cost of error correction is set to  $\lambda_{\text{EC}}(\epsilon_c) = \log_2(1/\epsilon_c)$ . We also assume  $n_Z = n_{\text{rep}}\tilde{p}_Z^2$  and  $n_X = n_{\text{rep}}\tilde{p}_X^2$ .

If we do not care about the key length for  $k_X > 0$ , the optimal choice of  $f(k_X, n_X, n_{\text{tot}})$  satisfying Eq. (33) (or Eq. (32)) is given by  $f(k_X, n_X, n_{\text{tot}}) = n_{\text{tot}} - n_X$  for  $k_X \geq 1$  and  $f(0, n_X, n_{\text{tot}}) = f_{\text{opt}}^{(k_X=0)}(n_X, n_{\text{tot}})$  with

$$f_{\text{opt}}^{(k_X=0)}(n_X, n_{\text{tot}}) := \min \left\{ k_{\text{tot}} \mid G(n_X; k_{\text{tot}}, n_{\text{tot}}) \leq \epsilon_{\text{PE}} \right\} - 1$$

$$G(n_X; k_{\text{tot}}, n_{\text{tot}}) := \sum_{n_X \leq n'_X \leq n_{\text{tot}} - k_{\text{tot}}} \text{HG}(0; n'_X, k_{\text{tot}}, n_{\text{tot}}) \text{BI}(n'_X; n_{\text{tot}}, p_X). \quad (34)$$

The proof is analogous to the one for  $f_{\text{BI}}(k_X)$  or  $f_{\text{HG}}(k_X, n_X, n_{\text{tot}})$ . Since  $G(n_X; k_{\text{tot}}, n_{\text{tot}})$  is a decreasing function of  $k_{\text{tot}}$ , by using an argument similar to the one leading to Eq. (21), we have

$$\sum_{n_X; k_{\text{tot}} > f_{\text{opt}}^{(k_X=0)}(n_X, n_{\text{tot}})} \text{HG}(0; n_X, k_{\text{tot}}, n_{\text{tot}}) \text{BI}(n_X; n_{\text{tot}}, p_X) \leq \epsilon_{\text{PE}}. \quad (35)$$

This is identical to Eq. (33) since  $k_X < k_{\text{tot}} - f(k_X, n_X, n_{\text{tot}})$  is never satisfied for  $k_X \geq 1$ . The key length when  $k_X = 0$  was observed is then given by

$$l^{(\text{opt})} := n_Z \left( 1 - h \left( \frac{f_{\text{opt}}^{(k_X=0)}(n_X, n_{\text{tot}})}{n_Z} \right) \right) - \log_2 \frac{2}{\epsilon_{\text{PA}}} - \lambda_{\text{EC}}(\epsilon_c). \quad (36)$$

In Fig. 1, we show the secure key ratios to the asymptotic case  $l^{(\text{BI})}/n_{\text{rep}}$  (red, bottom),  $l^{(\text{HG})}/n_{\text{rep}}$  (blue, middle) and  $l^{(\text{opt})}/n_{\text{rep}}$  (dashed, top) as functions of total rounds of the protocol  $n_{\text{rep}}$ . For each  $n_{\text{rep}}$ , the value of  $\tilde{p}_X$  was optimized to

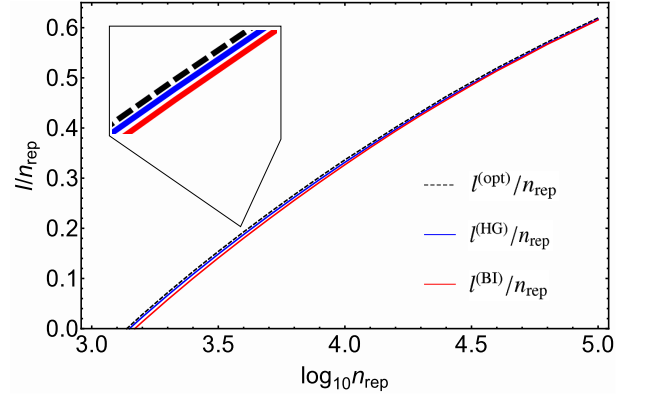


FIG. 1. Secure key ratio of the qubit-based BB84 protocol to the asymptotic limit as a function of total rounds of the protocol  $n_{\text{rep}}$ . We assume no errors ( $k_X = 0$ ) and no loss ( $n_{\text{tot}} = n_{\text{rep}}$ ). The security parameters are set to  $\epsilon_c = 10^{-15}$  and  $\epsilon_s = 10^{-10}$ . The top, middle and bottom curves represent the ratios  $l^{(\text{opt})}/n_{\text{rep}}$ ,  $l^{(\text{HG})}/n_{\text{rep}}$  (method with simple random sampling) and  $l^{(\text{BI})}/n_{\text{rep}}$  (Bernoulli-sampling method), respectively. In the limit of  $n_{\text{rep}} \rightarrow \infty$ , each curve converges to  $l/n_{\text{rep}} = 1$ .

maximize the key length. In the limit of  $n_{\text{rep}} \rightarrow \infty$ , each curve converges to  $l/n_{\text{rep}} = 1$ . The security parameters are set to  $\epsilon_c = 10^{-15}$  and  $\epsilon_s = 10^{-10}$ ,  $\epsilon_{\text{PE}} = 1/4 \times 10^{-20}$  and  $\epsilon_{\text{PA}} = 1/4 \times 10^{-20}$ . We see that although the key rate  $l^{(\text{opt})}$  is the best, the three methods achieve almost the same key length.

## IV. ANALYSIS FOR WCP-BASED PROTOCOL

Here, we apply the analyses introduced in the previous section to the protocols using weak coherent pulses (WCP). We consider the WCP-based BB84 protocol in the subsections A and B, and move to the DQPS protocol in subsection C.

### A. The WCP-BB84 protocol

Similarly to the ideal qubit-based protocol, the WCP-BB84 protocol also follows the procedures described in Sec. II, but the latter assumes more general light sources and measurement apparatuses. We prove the security of the WCP-BB84 protocol based on that of qubit-based BB84 protocol combined with GLLP's tagging idea [20]. We impose the following assumptions on Alice and Bob's devices. For Alice's light source, we assume that the four states  $\hat{\rho}_{W,a}$  with  $W \in \{Z, X\}$  and  $a \in \{0, 1\}$  are written as

$$\hat{\rho}_{W,a} = (1 - r_{\text{tag}})\hat{\rho}_{W,a,\text{unt}} \oplus r_{\text{tag}}\hat{\rho}_{W,a,\text{tag}}. \quad (37)$$

For  $\{\hat{\rho}_{W,a,\text{unt}}\}$ , we assume that there is a basis-independent state  $\hat{\chi}_{\text{unt}}$  on the system AS satisfying

$$\text{tr}_A \left( (|a_W\rangle\langle a_W|_A \otimes \hat{\mathbb{1}}_S) \hat{\chi}_{\text{unt}} \right) = \frac{1}{2} \hat{\rho}_{W,a,\text{unt}}. \quad (38)$$

We place no restriction on the states  $\{\hat{\rho}_{W,a,\text{tag}}\}$ . We may still find a state  $\hat{\chi}_{W,\text{tag}}$  on the system  $AS$  such that

$$\text{tr}_A \left( (|a_W\rangle\langle a_W|_A \otimes \mathbb{1}_S) \hat{\chi}_{W,\text{tag}} \right) = \frac{1}{2} \hat{\rho}_{W,a,\text{tag}}, \quad (39)$$

but  $\hat{\chi}_{W,\text{tag}}$  depends on a selected basis  $W \in \{Z, X\}$ . The form of Eq. (37) allows an interpretation that each round is classified as either tagged or untagged [20]. The density operator  $\hat{\rho}_{W,a,\text{unt}}$  is the state of Alice's untagged signal which may generate a secure key and  $\hat{\rho}_{W,a,\text{tag}}$  is that of her tagged signal which is considered to be totally insecure. Equation (38) indicates that Alice's basis choice can be postponed until Bob receives the system  $S$  for untagged rounds. Eqs. (37) and (38) are realized, for example, if Alice uses a laser emitting an ideally polarized coherent pulse with mean photon number  $\mu$  and randomizes its optical phase. In this case,  $\hat{\rho}_{W,a,\text{unt}}$  and  $\hat{\rho}_{W,a,\text{tag}}$  are written as

$$(1 - r_{\text{tag}}) \hat{\rho}_{W,a,\text{unt}} = e^{-\mu} |0_{W,a}\rangle\langle 0_{W,a}| + \mu e^{-\mu} |1_{W,a}\rangle\langle 1_{W,a}| \quad (40)$$

$$r_{\text{tag}} \hat{\rho}_{W,a,\text{tag}} = e^{-\mu} \sum_{m=2}^{\infty} \frac{\mu^m}{m!} |m_{W,a}\rangle\langle m_{W,a}|, \quad (41)$$

where  $|m_{W,a}\rangle$  is an  $m$ -photon state with a basis  $W$  and a bit  $a$ , and

$$r_{\text{tag}} = 1 - e^{-\mu} - \mu e^{-\mu} \quad (42)$$

represents the probability that Alice emits two or more photons. Our proof does not depend on the specific model such as coherent light source, but depends only on Eqs. (37) and (38).

For Bob's measurement apparatus, we impose either of the following two assumptions.

- (i) The probability of detecting a signal at Bob's receiver is independent of his basis choice.
- (ii) The measurement of an input signal on the system  $S$  is replaced by an ideal single-photon measurement on the system  $B$  preceding by a squashing operation [21, 22].

The condition (i), which is weaker than condition (ii), allows us to use the security proof with complementarity [23] and uncertainty principle [4]. The condition (ii) validates the use of the security proof with entanglement distillation [17]. For the WCP-BB84 protocol, both conditions are satisfied if we assume the following model for Bob's apparatus: Bob actively chooses the basis, and uses two threshold detectors corresponding to the measurement result "0" and "1" after a polarization beam splitter. He assigns random bit if both detectors report their detections. In addition, the inefficiency and dark countings of the detectors are allowed as long as they are equivalently represented by an absorber and a stray photon source placed in front of Bob's apparatus.

For the proof with entanglement distillation, we use Eqs. (38), (39) and the assumption (ii) of Bob's apparatus to convert the actual protocol equivalently to a protocol in which Alice and Bob make ideal measurements on the qubit systems  $A$  and  $B$ . Then a phase error is defined in the same way to the ideal BB84 protocol, namely, an error between Alice's and Bob's outcomes of ideal  $X$ -basis measurements ( $|0_X\rangle, |1_X\rangle$ ) on a  $Z$ -labeled round. As for the proof with complementarity, phase error in a  $Z$ -labeled round is defined as an error occurring when Alice makes an ideal  $X$ -basis measurement on

the system  $A$  and Bob makes the actual  $X$ -basis measurement on the system  $S$  (the measurement conducted on  $X$ -labeled rounds in the actual protocol).

The secure key length formulated in Eqs. (9)-(12) for the ideal protocol can be adapted to the WCP protocol through the tagging idea. Let  $k_{\text{ph,unt}}$  be the total number of phase errors on the untagged  $Z$ -labeled rounds. Let  $n_{Z,\text{unt}}$  be the number of untagged  $Z$ -labeled rounds. Since each round can be classified as tagged or untagged,  $n_{Z,\text{unt}}$  is a well-defined random variable in the actual protocol. Suppose that an upper bound of  $k_{\text{ph,unt}}$  is given as a function of  $k_X, n_X, n_{\text{tot}}$  and  $n_{Z,\text{unt}}$ :

$$\Pr(k_{\text{ph,unt}} > f(k_X, n_X, n_{\text{tot}}, n_{Z,\text{unt}})) \leq \epsilon_{\text{PE}}. \quad (43)$$

According to the tagging idea, the final key is  $\epsilon_s$ -secret if

$$n_{\text{fin}} \leq n_{Z,\text{unt}} \left( 1 - h \left( \frac{f(k_X, n_X, n_{\text{tot}}, n_{Z,\text{unt}})}{n_{Z,\text{unt}}} \right) \right) - \log_2 \frac{2}{\epsilon_{\text{PA}}} - \lambda_{\text{EC}}(\epsilon_c) \quad (44)$$

is satisfied. In the practical situation, the exact value of  $n_{Z,\text{unt}}$  is not available, and hence it is impossible to satisfy Eq. (44) with certainty. Instead, we allow a small error probability  $\epsilon_{Z,\text{unt}}$ . Suppose that there is a probabilistic lower bound  $\underline{n}_{Z,\text{unt}}$  which satisfies

$$\Pr(n_{Z,\text{unt}} < \underline{n}_{Z,\text{unt}}) \leq \epsilon_{Z,\text{unt}}. \quad (45)$$

A key length as a function of observed values  $k_X, n_X, n_{\text{tot}}$  is given by minimizing the right-hand side of Eq. (44) in the range of  $n_{Z,\text{unt}} \geq \underline{n}_{Z,\text{unt}}$ .

Under the assumptions for the source and measurement apparatus, the basic distributions used in the previous section, Eqs. (13) and (14), are still valid if we confine ourselves to the untagged rounds. Although the fact may be intuitively obvious for the WCP-BB84 protocol, here we give its mathematical justification since it helps when we treat a less intuitive protocol in Subsection IV C. We define a set of integers labeling the rounds in the protocol as  $\mathcal{N}_{\text{rep}} := \{1, 2, \dots, n_{\text{rep}}\}$ . As subsets of  $\mathcal{N}_{\text{rep}}$ , let us define the set of the integers labeling the rounds where Alice (Bob) chooses  $X$  basis as  $\mathcal{X}_A$  ( $\mathcal{X}_B$ ) regardless of detection. Define those labeling the untagged and detected rounds as  $\mathcal{N}_{\text{unt}}$ . Let  $\mathcal{K}_{\text{unt}}$  be a subset of  $\mathcal{N}_{\text{unt}}$  labeling the rounds which have errors when Alice and Bob conduct virtual  $X$ -basis measurements regardless of their basis choice. For any subset  $\mathcal{M}$ , let  $\overline{\mathcal{M}} := \mathcal{N}_{\text{rep}} \setminus \mathcal{M}$ . With these notations,  $k_{\text{ph,unt}} = |\overline{\mathcal{X}}_A \cap \overline{\mathcal{X}}_B \cap \mathcal{K}_{\text{unt}}|$  and  $n_{Z,\text{unt}} = |\overline{\mathcal{X}}_A \cap \overline{\mathcal{X}}_B \cap \mathcal{N}_{\text{unt}}|$ . We define other random variables as follows:  $k_{X,\text{unt}} := |\mathcal{X}_A \cap \mathcal{X}_B \cap \mathcal{K}_{\text{unt}}|$ ,  $n_{X,\text{unt}} := |\mathcal{X}_A \cap \mathcal{X}_B \cap \mathcal{N}_{\text{unt}}|$ ,  $k_{\text{tot,unt}} := k_{X,\text{unt}} + k_{\text{ph,unt}}$  and  $n_{\text{tot,unt}} := n_{X,\text{unt}} + n_{Z,\text{unt}}$ . From the assumption of Alice's source Eq. (38) and the assumptions of Bob's receiver (i) or (ii), the choice of  $\mathcal{X}_A \cap \mathcal{N}_{\text{unt}}$  and  $\mathcal{X}_B \cap \mathcal{N}_{\text{unt}}$  can be postponed after  $\mathcal{N}_{\text{unt}}$  and  $\mathcal{K}_{\text{unt}}$  are determined as far as untagged incidents are concerned. Then we have

$$\begin{aligned} \Pr(\mathcal{X}_A \cap \mathcal{N}_{\text{unt}} = \mathcal{M}_A, \mathcal{X}_B \cap \mathcal{N}_{\text{unt}} = \mathcal{M}_B \mid \mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}) \\ = \Theta(\mathcal{M}_A, \mathcal{N}_{\text{unt}}) \Theta(\mathcal{M}_B, \mathcal{N}_{\text{unt}}) \end{aligned} \quad (46)$$

for all  $\mathcal{M}_A \subset \mathcal{N}_{\text{unt}}$  and  $\mathcal{M}_B \subset \mathcal{N}_{\text{unt}}$ , where we defined

$$\Theta(\mathcal{M}_1, \mathcal{M}_2) = \tilde{p}_X^{|\mathcal{M}_1|} \tilde{p}_Z^{|\mathcal{M}_2 \setminus \mathcal{M}_1|}. \quad (47)$$

By simple calculation of the probability theory, we have

$$\Pr(k_{X,\text{unt}} | k_{\text{tot,unt}}, n_{\text{tot,unt}}) = \text{BI}(k_{X,\text{unt}}; k_{\text{tot,unt}}, p_X) \quad (48)$$

and

$$\Pr(k_{X,\text{unt}} | n_{X,\text{unt}}, k_{\text{tot,unt}}, n_{\text{tot,unt}}) = \text{HG}(k_{X,\text{unt}}; n_{X,\text{unt}}, k_{\text{tot,unt}}, n_{\text{tot,unt}}), \quad (49)$$

which means that Eqs. (13) and (14) essentially hold true for the untagged rounds.

Now we derive a key rate formula for the WCP BB84 protocol based on Eq. (48), as was done with the Bernoulli-sampling method for the qubit-based protocol in Sec. III B. First, we seek for  $f(k_X, n_X, n_{\text{tot}}, n_{Z,\text{unt}})$  which satisfies Eq. (43). Analogous to the derivation of Eq. (23) from Eq. (13), Eq. (48) leads to

$$\Pr(k_{\text{ph,unt}} > f_{\text{BI}}(k_{X,\text{unt}}) | k_{\text{tot,unt}}, n_{\text{tot,unt}}) \leq \epsilon_{\text{PE}} \quad (50)$$

for any  $k_{\text{tot,unt}}$  and  $n_{\text{tot,unt}}$ , and hence we have

$$\Pr(k_{\text{ph,unt}} > f_{\text{BI}}(k_{X,\text{unt}})) \leq \epsilon_{\text{PE}}. \quad (51)$$

Since  $k_{X,\text{unt}}$  is not an observed value, we use the obvious bound

$$k_{X,\text{unt}} \leq k_X. \quad (52)$$

Using the inequality

$$\begin{aligned} C_{\text{BI}}(k_X + 1; k_{\text{tot}} + 1, p_X) \\ = C_{\text{BI}}(k_X; k_{\text{tot}}, p_X) + (1 - p_X)\text{BI}(k_X + 1; k_{\text{tot}}, p_X) \\ \geq C_{\text{BI}}(k_X; k_{\text{tot}}, p_X) \end{aligned} \quad (53)$$

in Eq. (18), we have  $f_{\text{BI}}(k_X) \leq f_{\text{BI}}(k_X + 1)$ , implying that  $f_{\text{BI}}(k_X)$  is an increasing function. Hence, Eqs. (51) and (52) lead to

$$\Pr(k_{\text{ph,unt}} > f_{\text{BI}}(k_X)) \leq \epsilon_{\text{PE}}, \quad (54)$$

which means that  $f_{\text{BI}}(k_X)$  fulfills Eq. (43).

Next, we determine  $\underline{n}_{Z,\text{unt}}$  which satisfies Eq. (45). To determine a lower bound of  $n_{Z,\text{unt}}$ , we consider an upper bound of  $n_{Z,\text{tag}} := n_Z - n_{Z,\text{unt}}$ . Let  $N_{Z,\text{tag}}$  be the number of rounds where Alice chooses  $Z$  basis, Bob chooses  $Z$  basis and the light source emits a tagged signal (two photons or more). As those conditions are independent of each other as seen from Eq. (37), we have

$$\Pr(N_{Z,\text{tag}}) = \text{BI}(N_{Z,\text{tag}}, n_{\text{rep}}, r_{\text{tag}} \tilde{p}_Z^2). \quad (55)$$

Since  $n_{Z,\text{tag}}$  is the number of detected rounds among the  $N_{Z,\text{tag}}$  rounds,

$$n_{Z,\text{tag}} \leq N_{Z,\text{tag}} \quad (56)$$

holds. Eqs. (55) and (56) lead to

$$\Pr(n_{Z,\text{tag}} > n) \leq 1 - C_{\text{BI}}(n; n_{\text{rep}}, r_{\text{tag}} \tilde{p}_Z^2) \quad (57)$$

for any  $n$ . Thus, we have

$$\Pr(n_{Z,\text{tag}} > g(r_{\text{tag}} \tilde{p}_Z^2, \epsilon_{Z,\text{unt}})) \leq \epsilon_{Z,\text{unt}}, \quad (58)$$

where

$$g(x, y) := \min \left\{ n \mid 1 - C_{\text{BI}}(n; n_{\text{rep}}, x) \leq y \right\}. \quad (59)$$

Let  $\underline{n}_{Z,\text{unt}}$  be

$$\underline{n}_{Z,\text{unt}} := n_Z - g(r_{\text{tag}} \tilde{p}_Z^2, \epsilon_{Z,\text{unt}}). \quad (60)$$

By using  $n_{Z,\text{tag}} = n_Z - n_{Z,\text{unt}}$ , Eq. (58) leads to

$$\Pr(n_{Z,\text{unt}} < \underline{n}_{Z,\text{unt}}) \leq \epsilon_{Z,\text{unt}}. \quad (61)$$

Since  $n_{Z,\text{unt}}$  is known in principle in the actual protocol, the final state  $\rho_{AE}^{\text{fin}}$  is written as a direct sum of the part for  $n_{Z,\text{unt}} < \underline{n}_{Z,\text{unt}}$  and the one for  $n_{Z,\text{unt}} \geq \underline{n}_{Z,\text{unt}}$ . Hence, combined with Eqs. (44), (54) and (61), by setting

$$\epsilon_s = \sqrt{2} \sqrt{\epsilon_{\text{PE}} + \epsilon_{\text{PA}}} + \epsilon_{Z,\text{unt}}, \quad (62)$$

the protocol is  $\epsilon_c$ -correct and  $\epsilon_s$ -secret if

$$n_{\text{fin}} \leq l_{\text{WCP}}^{(\text{BI})} := \underline{n}_{Z,\text{unt}} \left( 1 - h \left( \frac{f_{\text{BI}}(k_X)}{\underline{n}_{Z,\text{unt}}} \right) \right) - \log_2 \frac{2}{\epsilon_{\text{PA}}} - \lambda_{\text{EC}}(\epsilon_c). \quad (63)$$

Together with Eqs. (18), (19), (59) and (60), Eq. (63) constitutes the main result of Sec IV A.

For the purpose of comparison, here we also discuss what the key rate formula looks like if we start from Eq. (49), based on simple random sampling. As we have derived Eq. (28) from Eq. (14), Eq. (49) leads to

$$\Pr(k_{\text{ph,unt}} > f_{\text{HG}}(k_{X,\text{unt}}, n_{X,\text{unt}}, n_{\text{tot,unt}}) | k_{\text{tot,unt}}, n_{\text{tot,unt}}) \leq \epsilon_{\text{PE}}, \quad (64)$$

which, in turn, leads to

$$\Pr(k_{\text{ph,unt}} > f_{\text{HG}}(k_{X,\text{unt}}, n_{X,\text{unt}}, n_{\text{tot,unt}})) \leq \epsilon_{\text{PE}}. \quad (65)$$

Similarly to  $f_{\text{BI}}(k_X)$ , we can prove that  $f_{\text{HG}}(k_X, n_X, n_{\text{tot}})$  is an increasing function of  $k_X$ . Since  $k_{X,\text{unt}}$  is upper-bounded by Eq. (52), Eq. (65) leads to

$$\Pr(k_{\text{ph,unt}} > f_{\text{HG}}(k_X, n_{X,\text{unt}}, n_{\text{tot,unt}})) \leq \epsilon_{\text{PE}}. \quad (66)$$

In contrast to Eq. (54), it requires an additional estimation process for  $n_{X,\text{unt}}$  to obtain  $f_{\text{HG}}(k_X, n_{X,\text{unt}}, n_{\text{tot,unt}})$ . A lower bound defined by  $\underline{n}_{X,\text{unt}} := n_X - g(r_{\text{tag}} \tilde{p}_X^2, \epsilon_{X,\text{unt}})$  satisfies

$$\Pr(n_{X,\text{unt}} < \underline{n}_{X,\text{unt}}) \leq \epsilon_{X,\text{unt}}. \quad (67)$$

Combined with Eqs. (44), (61), (66), and (67), by setting

$$\epsilon_s = \sqrt{2} \sqrt{\epsilon_{\text{PE}} + \epsilon_{\text{PA}}} + \epsilon_{Z,\text{unt}} + \epsilon_{X,\text{unt}}, \quad (68)$$

the protocol is  $\epsilon_c$ -correct and  $\epsilon_s$ -secret if

$$\begin{aligned} n_{\text{fin}} \leq l_{\text{WCP}}^{(\text{HG})} &:= \min_{n_{Z,\text{unt}} \geq \underline{n}_{Z,\text{unt}}} \xi(k_X, \underline{n}_{X,\text{unt}}, n_{Z,\text{unt}}) \\ \xi(k_X, \underline{n}_{X,\text{unt}}, n_{Z,\text{unt}}) &:= \tilde{\xi}(k_X, \underline{n}_{X,\text{unt}}, n_{Z,\text{unt}}) - \log_2 \frac{2}{\epsilon_{\text{PA}}} - \lambda_{\text{EC}}(\epsilon_c) \\ \tilde{\xi}(k_X, \underline{n}_{X,\text{unt}}, n_{Z,\text{unt}}) &:= n_{Z,\text{unt}} \left( 1 - h \left( \frac{f_{\text{HG}}(k_X, \underline{n}_{X,\text{unt}}, \underline{n}_{X,\text{unt}} + n_{Z,\text{unt}})}{n_{Z,\text{unt}}} \right) \right). \end{aligned} \quad (69)$$



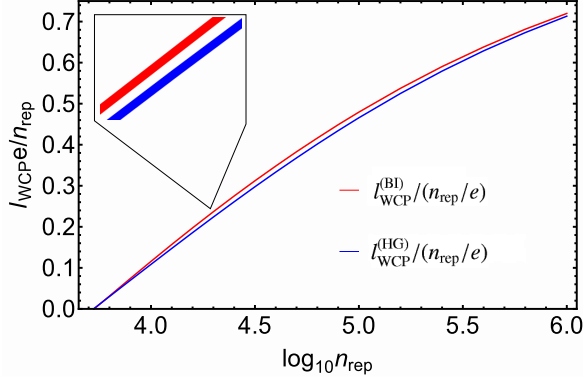


FIG. 2. Comparison of estimation methods for the WCP-BB84 protocol. Upper curve (Bernoulli-sampling method): Secure key ratio to the asymptotic limit  $l_{\text{WCP}}^{(\text{BI})}/(n_{\text{rep}}/e)$  as a function of total rounds of the protocol  $n_{\text{rep}}$ . Lower curve (method with simple random sampling): An upper bound on the derived secure key ratio  $l_{\text{WCP}}^{(\text{HG})}/(n_{\text{rep}}/e)$ . We assume no error ( $k_X = 0$ ) and no loss ( $n_{\text{tot}} = n_{\text{rep}}(1 - e^{-\mu})$ ). The security parameters are set to  $\epsilon_c = 10^{-15}$  and  $\epsilon_s = 10^{-10}$ . In the limit of  $n_{\text{rep}} \rightarrow \infty$ , each curve converges to  $l_{\text{WCP}}/(n_{\text{rep}}/e) = 1$ .

The reason that the minimization of  $n_{Z,\text{unt}}$  appears is because  $\tilde{\xi}(k_X, \underline{n}_{X,\text{unt}}, n_{Z,\text{unt}})$  is not monotone-increasing function of  $n_{Z,\text{unt}}$ . For example, with  $\epsilon_{\text{PE}} = 1/16 \times 10^{-20}$ , we numerically confirmed that  $\tilde{\xi}(0, 25000, 25318) \sim 24631$  and  $\tilde{\xi}(0, 25000, 25319) \sim 24623$ . This means that the protocol with final key length  $l = \xi(k_X, \underline{n}_{X,\text{unt}}, n_{Z,\text{unt}})$  is not necessarily secure.

As can be seen from the comparison between Eqs. (63) and (69), the method with simple random sampling is much more complicated than the Bernoulli-sampling method, involving an additional estimated parameter and a minimization. Moreover, as shown in Sec. IV B, it tends to give a key rate lower than the Bernoulli-sampling method, probably because of the use of pessimistic bound on  $n_{X,\text{unt}}$ .

### B. Numerical examples

Here, we show two examples of numerical calculation for the WCP-BB84 protocol. We assume that the light source emits a pulse whose photon-number distribution is Poissonian with mean  $\mu$ , namely, Eq. (42). Like Fig. 1 for the ideal protocol, we first calculated the simplest case where no error is observed ( $k_X = 0$ ) and no loss occurs ( $n_{\text{tot}} = n_{\text{rep}}(1 - e^{-\mu})$ ), which is shown in Fig. 2. The cost of error correction was set to  $\lambda_{\text{EC}}(\epsilon_c) = \log_2(1/\epsilon_c)$ . We assumed  $n_Z = n_{\text{tot}}\tilde{p}_Z^2$  and  $n_X = n_{\text{tot}}\tilde{p}_X^2$ . The values of  $\tilde{p}_X$  and  $\mu$  were optimized for each value of  $n_{\text{rep}}$ . For calculation of  $l_{\text{WCP}}^{(\text{BI})}$ , the security parameters were set to  $\epsilon_c = 10^{-15}$ ,  $\epsilon_s = 10^{-10}$ ,  $\epsilon_{\text{PE}} = 1/16 \times 10^{-20}$ ,  $\epsilon_{\text{PA}} = 1/16 \times 10^{-20}$  and  $\epsilon_{Z,\text{unt}} = 1/2 \times 10^{-10}$ . The result is shown as the red curve in Fig. 2, where the key length Eq. (63) is normalized by the optimized asymptotic key rate of  $1/e$  per signal at  $\mu = 1$  and  $\tilde{p}_X \rightarrow 0$ . We see that a final key can be extracted when the total rounds  $n_{\text{rep}}$  is more than  $\sim 10^{3.7}$ .

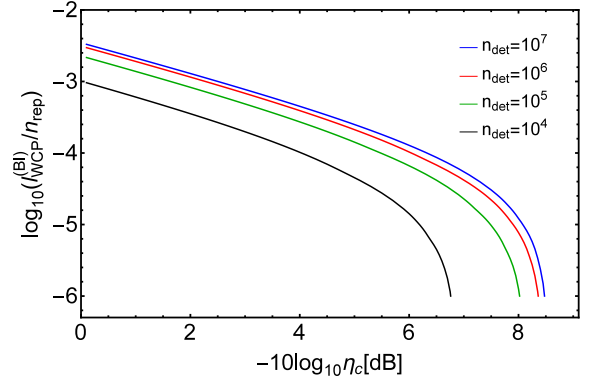


FIG. 3. Secure key rate per signal of the WCP-BB84 protocol  $l_{\text{WCP}}^{(\text{BI})}/n_{\text{rep}}$  as a function of channel transmission  $\eta_c$ . The parameters are set to be the same as Ref. [24]. Quantum efficiency of detectors:  $\eta_d = 0.1$ . Dark count probability per pulse per detector:  $p_{\text{dark}} = 10^{-5}$ . Loss-independent bit error: 0.5%. Error correction cost:  $\lambda_{\text{EC}}(\epsilon_c) = 1.05h(E/Q) + \log_2(1/\epsilon_c)$ . The security parameters:  $\epsilon_c = 10^{-10}$  and  $\epsilon_s = 10^{-5}$ . From the top to the bottom curve, the number of detected signals are  $n_{\text{det}} = 10^7, 10^6, 10^5$  and  $10^4$ , respectively. The required number of detected signals to generate a final key is less than  $10^4$ , while it was  $\sim 10^7$  in the previous result [24].

while the threshold is  $n_{\text{rep}} \sim 10^{3.2}$  for the ideal protocol using the same parameters (see also Fig. 1). For comparison, we also calculated the value of  $\xi(k_X, \underline{n}_{X,\text{unt}}, n_{Z,\text{unt}})/(n_{\text{rep}}/e)$  under the same condition, which is shown as the blue (lower) curve in Fig. 2. The security parameters were the same as the red (upper) curve, except for  $\epsilon_{Z,\text{unt}} = \epsilon_{X,\text{unt}} = 1/4 \times 10^{-10}$ . The quantity  $\xi(k_X, \underline{n}_{X,\text{unt}}, n_{Z,\text{unt}})$  is an upper bound of  $l_{\text{WCP}}^{(\text{HG})}$  derived in Eq. (69). The figure shows that the key length  $l_{\text{WCP}}^{(\text{BI})}$  from Bernoulli sampling is higher than  $l_{\text{WCP}}^{(\text{HG})}$  from simple random sampling. A possible reason is that the estimation of  $\underline{n}_{X,\text{unt}}$ , which is a pessimistic bound of  $n_{X,\text{unt}}$ , is not required in determining  $f_{\text{BI}}(k_X)$ .

In Fig. 3, we show a result in more practical situations based on Eq. (63) to make comparison to the previous finite-key analysis for the WCP-BB84 protocol [24]. The figure shows the dependence of secure key rate  $l_{\text{WCP}}^{(\text{BI})}/n_{\text{rep}}$  on the channel transmission  $\eta_c$ . In each curve, the number of Bob's detected signals  $n_{\text{det}}$  is fixed as  $n_{\text{det}} = 10^4, 10^5, 10^6$  and  $10^7$  (from the bottom to the top, respectively). The parameters were chosen to be the same as [24]: Quantum efficiency of both detectors is  $\eta_d = 0.1$  and a dark count probability per pulse is  $p_{\text{dark}} = 10^{-5}$  per detector. In addition to errors from dark counts, there is a 0.5% loss-independent bit error. The security parameters were set to  $\epsilon_c = 10^{-10}$ ,  $\epsilon_s = 10^{-5}$ ,  $\epsilon_{\text{PE}} = 1/16 \times 10^{-10}$ ,  $\epsilon_{\text{PA}} = 1/16 \times 10^{-10}$ , and  $\epsilon_{Z,\text{unt}} = 1/2 \times 10^{-5}$ . Total transmission rate is  $Q = 1 - (1 - 2p_{\text{dark}})e^{-\mu\eta_c\eta_d}$ , and error rate is given by  $E/Q$  where  $E = 0.005(1 - e^{-\mu\eta_c\eta_d}) + p_{\text{dark}}e^{-\mu\eta_c\eta_d}$ . Based on the parameters above, we assume  $\lambda_{\text{EC}}(\epsilon_c) = 1.05h(E/Q) + \log_2(1/\epsilon_c)$ ,  $n_{\text{rep}} = n_{\text{det}}/Q$ ,  $n_Z = n_{\text{det}}\tilde{p}_Z^2$ ,  $n_X = n_{\text{det}}\tilde{p}_X^2$  and  $k_X = n_X E/Q$ . To save the computation time, we used Chernoff

bound [25]

$$C_{\text{BI}}(k_X; k_{\text{tot}}, p_X) \leq D\left(\frac{k_X}{k_{\text{tot}}}, k_{\text{tot}}, p_X\right) \quad (70)$$

for  $(k_X, k_{\text{tot}}, p_X)$  satisfying  $k_X \leq k_{\text{tot}} p_X$ , where

$$D(x, y, z) := \left( \left( \frac{z}{x} \right)^x \left( \frac{1-z}{1-x} \right)^{1-x} \right)^y. \quad (71)$$

In Fig. 3, we see that a key can be extracted even when  $n_{\text{det}} = 10^4$ . This is a significant improvement from the result of [24], in which the required number of detected signals to generate a final key is  $n_{\text{det}} \sim 10^7$ .

### C. The DQPS protocol

In this section, we conduct finite-key analysis of the DQPS protocol based on the property of binomial distribution Eq. (48). The security of the DQPS protocol was recently proved in the asymptotic limit [13]. The DQPS protocol uses encoding on four relative phases  $\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$  between neighboring pulses in a pulse train of fixed length  $L$ . The DQPS protocol has essentially the same setup as the BB84 protocol with phase encoding (PE-BB84 protocol), which can be regarded as the DQPS protocol with  $L = 2$ . In Ref. [13], we showed that the secure key rate of the DQPS protocol is  $8/3$  as high as that of the PE-BB84 protocol in the asymptotic limit. However, since the security proof is not so straightforward as that of the BB84 protocol, it is not trivial whether the advantage of the DQPS protocol over the PE-BB84 protocol still holds considering the statistical fluctuations in the finite-key case. This motivates us to conduct finite-key analysis for the DQPS protocol by using the Bernoulli-sampling method proposed in this work.

The overview of the DQPS protocol is shown in Fig. 4. The precise description of the protocol and physical assumptions for the security proof is given in Appendix. In the protocol,  $Z$  basis is chosen with probability  $\tilde{p}_Z$  to generate keys and  $X$  basis is chosen with probability  $\tilde{p}_X$  for leak monitoring. Relative phases between adjacent pulses are modulated by  $\{0, \pi\}$  for  $Z$  basis and  $\{\frac{\pi}{2}, \frac{3\pi}{2}\}$  for  $X$  basis. The protocol regards  $L$ -successive pulses as a block, and at most one key bit is extracted from each block. The randomization of the optical phase is conducted to the whole block, and a basis is also chosen for each block. Bob's receiver is composed of delayed interferometer with its delay being equal to the interval  $\Delta\tau$  of adjacent pulses. The longer arm of the interferometer incurs phase shift  $0$  ( $Z$ -basis) or  $\pi/2$  ( $X$ -basis), which are chosen with probability  $\tilde{p}_Z$  and  $\tilde{p}_X$ , respectively. After the interferometer, the pulses are measured by two photon detectors corresponding to bit values "0" and "1". If there is a detection from the superposition of the  $l$ -th and the  $(l-1)$ -th original pulses, we call it valid detection at  $l$ -th timing ( $1 \leq l \leq L-1$ ). An interference between different blocks at Bob's receiver is invalid and does not contribute to a key, which means that  $1/L$  of the whole detection events must be discarded. This is the origin

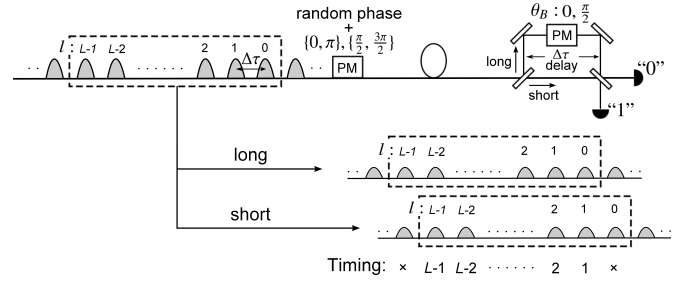


FIG. 4. Setup for the  $L$ -pulse DQPS protocol. The protocol regards a train of  $L$  pulses as a block, and the working basis is chosen for each block. At Alice's site, pulses are modulated with phase  $\{0, \pi, \frac{\pi}{2}, \frac{3\pi}{2}\}$  according to her random bit and basis choice. The randomization of the overall optical phase is also done for each block of  $L$  pulses. At Bob's site, each pulse train is fed to a delayed Mach-Zehnder interferometer with phase shift  $0$  or  $\frac{\pi}{2}$  according to his basis choice. Valid timings of detection are labeled by integers  $1, 2, \dots, L-1$ , according to the index of the pulse from the short arm of the interferometer. Detection from interference between pulses from different blocks is regarded as invalid and ignored.

of the advantage of the DQPS protocol over the PE-BB84 protocol, which is regarded as the DQPS protocol with  $L = 2$ .

In the DQPS protocol, application of the tagging idea is not straightforward since the chain of coherence among successive pulses prohibits us from defining the total photon number in neighboring two pulses. As a result, the conventional definition of tagging based on the emitted photon number cannot be applied here. In [13], we proposed an alternative approach to define the photon number indirectly through Bob's detection timing  $j$  and Alice's measurement result on her qubits, which enables us to assume that each round in the protocol is classified as either tagged or untagged. As a result, the variables  $k_{\text{ph,unt}}$  and  $n_{\text{Z,unt}}$  can be defined in the same way as in the WCP-BB84 protocol, and the argument up to Eq. (45) holds for the DQPS protocol as well. The remaining tasks are to find a function  $f$  satisfying Eq. (43) and to find a bound  $\underline{n}_{\text{Z,unt}}$  satisfying Eq. (45), both of which require slightly different approaches from the WCP-BB84 protocol.

Since our tagging definition for the DQPS protocol involves Bob's detection timing  $j$ , we cannot decompose the emitted states as in Eq. (37). Hence we need to justify Eq. (46) without using Eq. (38). This was essentially done in Ref. [13], which proved, in the notation of the present paper [26], that the joint probability of  $\mathcal{X}_A$ ,  $\mathcal{K}_{\text{unt}}$  and  $\mathcal{N}_{\text{unt}}$  is written in the following form:

$$\begin{aligned} \Pr(\mathcal{X}_A, \mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}) \\ = \Theta(\mathcal{X}_A, \mathcal{N}_{\text{rep}}) \beta'(\mathcal{X}_A \cap \overline{\mathcal{N}_{\text{unt}}}, \mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}). \end{aligned} \quad (72)$$

Since  $\Theta(\mathcal{M}, \mathcal{N}_{\text{rep}})$  defined in Eq. (47) satisfies

$$\Theta(\mathcal{M}, \mathcal{N}_{\text{rep}}) = \Theta(\mathcal{M} \cap \mathcal{N}_{\text{unt}}, \mathcal{N}_{\text{unt}}) \Theta(\mathcal{M} \cap \overline{\mathcal{N}_{\text{unt}}}, \overline{\mathcal{N}_{\text{unt}}}) \quad (73)$$

for any  $\mathcal{M} \subset \mathcal{N}_{\text{rep}}$ , from Eq. (72) we have

$$\begin{aligned} \Pr(\mathcal{X}_A \cap \mathcal{N}_{\text{unt}} = \mathcal{M}_A \mid \mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}) \\ = \Theta(\mathcal{M}_A, \mathcal{N}_{\text{unt}}) \gamma(\mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}) \end{aligned} \quad (74)$$

for any  $\mathcal{M}_A \subset \mathcal{N}_{\text{unt}}$ , where

$$\begin{aligned} \gamma(\mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}) \\ := \frac{\sum_{\mathcal{M}'_A \subset \mathcal{N}_{\text{unt}}} \Theta(\mathcal{M}'_A, \mathcal{N}_{\text{unt}}) \beta'(\mathcal{M}'_A, \mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}})}{\Pr(\mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}})}. \end{aligned} \quad (75)$$

Since the sum of  $\Theta(\mathcal{M}_A, \mathcal{N}_{\text{unt}})$  over  $\mathcal{M}_A$  is unity, Eq. (74) leads to  $\gamma(\mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}) = 1$ . Thus, we have

$$\begin{aligned} \Pr(\mathcal{X}_A \cap \mathcal{N}_{\text{unt}} = \mathcal{M}_A \mid \mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}) \\ = \Theta(\mathcal{M}_A, \mathcal{N}_{\text{unt}}). \end{aligned} \quad (76)$$

In the DQPS protocol, Bob's basis choice can be postponed after he confirms photon detection, which means that the choice of  $\mathcal{X}_B$  can be conducted after  $\mathcal{K}_{\text{unt}}$  and  $\mathcal{N}_{\text{unt}}$  are determined. Hence, we have

$$\begin{aligned} \Pr(\mathcal{X}_A \cap \mathcal{N}_{\text{unt}} = \mathcal{M}_A, \mathcal{X}_B \cap \mathcal{N}_{\text{unt}} = \mathcal{M}_B \mid \mathcal{K}_{\text{unt}}, \mathcal{N}_{\text{unt}}) \\ = \Theta(\mathcal{M}_A, \mathcal{N}_{\text{unt}}) \Theta(\mathcal{M}_B, \mathcal{N}_{\text{unt}}), \end{aligned} \quad (77)$$

which is identical to Eq. (46). Similarly to the WCP-BB84 protocol, Eq. (48) holds, which leads to Eq. (51):

$$\Pr(k_{\text{ph,unt}} > f_{\text{BI}}(k_X)) \leq \epsilon_{\text{PE}}. \quad (78)$$

The task of finding a bound  $\underline{n}_{\text{Z,unt}}$  satisfying Eq. (45) is done as follows. In Ref. [13], a modified protocol having exactly the same  $\Pr(n_{\text{Z,tag}})$  as the original protocol was introduced, in which a random variable  $N$  (denoted as  $n(c = d = 0, (z'_0 \dots z'_{L-1}) \notin \Gamma^{(m)})$  in Eq. (40) of Ref. [13]) satisfying  $N \geq n_{\text{Z,tag}}$  is defined. The variable obeys binomial distribution  $\text{BI}(N, n_{\text{rep}}, r_{\text{tag}} \tilde{p}_Z^2)$ , where  $r_{\text{tag}}$  is a property of the light source defined as

$$r_{\text{tag}} := 1 - \sum_m \text{tr}(\hat{\Pi}_S^{(m)} \hat{\sigma}_S), \quad (79)$$

where  $\hat{\sigma}_S$  is the state of  $L$  pulses emitted from Alice's light source, and  $\hat{\Pi}_S^{(m)}$  is a projector which is defined in Appendix. This implies that  $\Pr(n_{\text{Z,tag}})$  in the original protocol has the following property: There exists a function  $P(n_{\text{Z,tag}}, N)$  satisfying

$$\begin{aligned} \Pr(n_{\text{Z,tag}}) &= \sum_N P(n_{\text{Z,tag}}, N) \\ P(n_{\text{Z,tag}}, N) &= 0 \text{ for } n_{\text{Z,tag}} > N \\ \sum_{n_{\text{Z,tag}}} P(n_{\text{Z,tag}}, N) &= \text{BI}(N, n_{\text{rep}}, r_{\text{tag}} \tilde{p}_Z^2). \end{aligned} \quad (80)$$

This leads to

$$\Pr(n_{\text{Z,tag}} > n) \leq 1 - C_{\text{BI}}(n; n_{\text{rep}}, r_{\text{tag}} \tilde{p}_Z^2) \quad (81)$$

for any  $n$ , which is identical to Eq. (57). Then, following the same argument as the WCP-BB84 protocol, we see that

$$\Pr(n_{\text{Z,unt}} < \underline{n}_{\text{Z,unt}}) \leq \epsilon_{\text{Z,unt}} \quad (82)$$

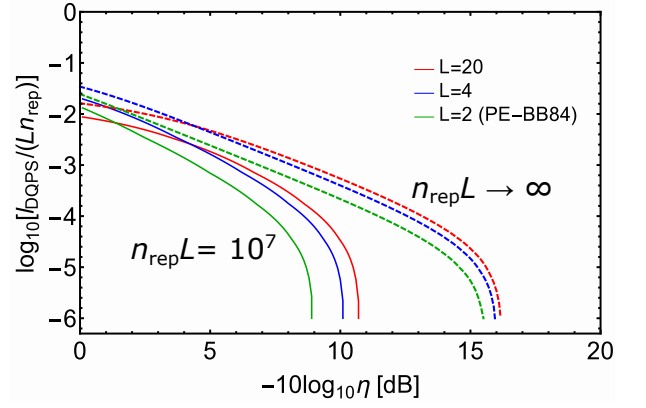


FIG. 5. Secure key rate per pulse of the DQPS protocol  $l_{\text{DQPS}}/(n_{\text{rep}}L)$  as a function of overall transmission  $\eta$ . Solid curves are the results of the finite key analysis with total pulse number  $n_{\text{rep}}L = 10^7$  and dashed curves are the results of the asymptotic case ( $n_{\text{rep}}L \rightarrow \infty$ ), which are obtained in Ref. [13]. For both solid and dotted curves, the top, middle and bottom curves represent the key rate for  $L = 20$ ,  $L = 4$  and  $L = 2$ , respectively. The parameters are set as follows. Dark count rate per pulse per detector:  $p_{\text{dark}} = 0.5 \times 10^{-5}$ . Loss-independent bit error: 3%. Cost for error correction:  $\lambda_{\text{EC}}(\epsilon_c) = 1.1h(E) + \log_2(1/\epsilon_c)$ . The security parameter:  $\epsilon_c = 10^{-15}$  and  $\epsilon_s = 10^{-10}$ . We see that the key rate of the DQPS protocol ( $L > 2$ ) is higher than that of the PE-BB84 protocol ( $L = 2$ ) for both the asymptotic and finite-key cases.

holds with

$$\underline{n}_{\text{Z,unt}} := n_Z - g(r_{\text{tag}} \tilde{p}_Z^2, \epsilon_{\text{Z,unt}}). \quad (83)$$

From Eqs. (44), (78) and (82), we arrive at a key rate formula which is identical to Eq. (63): The  $L$ -pulse DQPS protocol is  $\epsilon_c$ -correct and  $\epsilon_s$ -secret if the final key length  $n_{\text{fin}}$  satisfies

$$n_{\text{fin}} \leq l_{\text{DQPS}} := \underline{n}_{\text{Z,unt}} \left(1 - h\left(\frac{f_{\text{BI}}(k_X)}{\underline{n}_{\text{Z,unt}}}\right)\right) - \log_2 \frac{2}{\epsilon_{\text{PA}}} - \lambda_{\text{EC}}(\epsilon_c), \quad (84)$$

where  $\epsilon_s$  is given in Eq. (62). Together with Eqs. (18), (19), (59), (79) and (83), Eq. (84) constitutes the main result of Sec. IV C.

In Fig. 5, we show numerical results of secure key rate per pulse  $l_{\text{DQPS}}/(n_{\text{rep}}L)$  as a function of overall transmittance  $\eta := \eta_c \eta_d$  to compare the DQPS protocol ( $L > 2$ ) and the PE-BB84 protocol ( $L = 2$ ). The solid curves represent the key rate with fixed pulse number  $n_{\text{rep}}L = 10^7$ , and the dashed curves represent the one for the asymptotic case, which is obtained in our previous work [13]. We assumed that Alice generates a weak coherent pulse of mean photon number  $\mu$ . In this case,  $r_{\text{tag}}$  is given by

$$r_{\text{tag}} = 1 - \sum_{m=0}^{[L/2]} e^{-\mu L} \mu^m {}_{L+1-m}C_m. \quad (85)$$

Note that for  $L = 2$ ,  $r_{\text{tag}} = 1 - e^{-2\mu} - 2\mu e^{-2\mu}$  is identical to the probability that two or more photons are emitted in

a double-pulse signal in the PE-BB84 protocol. We assume dark count rate per pulse per detector  $p_{\text{dark}} = 0.5 \times 10^{-5}$  and a loss-independent bit error rate 3%. We also assumed that  $Q = 1 - (1 - 2(L-1)p_{\text{dark}})e^{-(L-1)\mu\eta}$ , reflecting the fact that there are  $L-1$  valid timings in a block. Error rate is given by  $E/Q$  where  $E = 0.03(1 - e^{-(L-1)\mu\eta}) + p_{\text{dark}}e^{-(L-1)\mu\eta}(L-1)$ . Based on these parameters, we assume  $\lambda_{\text{EC}}(\epsilon_c) = 1.1h(E/Q) + \log_2(1/\epsilon_c)$ ,  $n_Z = n_{\text{rep}}Q\tilde{p}_Z^2$ ,  $n_X = n_{\text{rep}}Q\tilde{p}_X^2$  and  $k_X = n_X E/Q$ . The values of  $\tilde{p}_X$  and  $\mu$  are optimized to maximize the key length. In the asymptotic limit, the parameter optimization leads to  $\tilde{p}_X \rightarrow 0$ ,  $n_{Z,\text{unt}} \rightarrow n_{\text{rep}}(Q - r_{\text{tag}})$  and  $f_{\text{BI}}(k_X)/n_{Z,\text{unt}} \rightarrow E/(Q - r_{\text{tag}})$  while  $Q$  and  $E$  are fixed. In finite-key cases, the Chernoff bound is used to calculate the key rate. The security parameters are set to be the same as those in Fig. 2. We see that the advantage of the DQPS protocol over the PE-BB84 protocol is maintained even if we include the effect of the finiteness of the key.

## V. CONCLUDING REMARKS

In this paper, we proposed a method of finite-key analysis based on Bernoulli sampling instead of simple random sampling. For the BB84 protocol using biased basis choice, the data gathered on one of the basis is solely used for estimation of the disturbance in the other basis, which enables us to regard the former as a sample drawn from the population via Bernoulli sampling. As a result, we obtained finite-sized key-length formulas based on the binomial distribution parametrized by the probability of the basis choice in the protocol. The appearance of the binomial distribution in our case is a direct consequence of the inherent statistics of the protocol, and it should be differentiated from the previous works which uses a binomial distribution to derive an upper bound on the hypergeometric distribution arising from simple random sampling.

The new method is particularly suited for the BB84 protocol with WCP. It enables simpler analysis compared to the method with simple random sampling since only the latter requires the estimation of the sample size ( $n_{X,\text{unt}}$ ). We may expect that this additional pessimistic bound makes the conventional method less efficient, which is corroborated by a numerical example showing that the key rate for the WCP-BB84 protocol obtained with our method is higher than that with simple random sampling. To make comparison with the previous finite-key analysis for the WCP-BB84 protocol [24], we calculated the key rate as a function of channel transmission and the number of detected signals, in the same practical parameter settings. The result shows that, while  $n_{\text{det}} \sim 10^7$  signals are necessary for producing a key in Ref. [24], our method only needs  $n_{\text{det}} \sim 10^4$  with the same parameters. In addition, the improved number  $10^4$  clarifies that the use of WCP instead of an ideal single photon causes only a small change in the finite-size effect. This was also confirmed in the numerical simulation assuming the perfect channel, in which the required number of rounds to generate a key is  $n_{\text{rep}} \sim 10^{3.7}$  for the WCP-BB84 protocol and is  $n_{\text{rep}} \sim 10^{3.2}$  for the single-photon BB84 protocol.

Finally, we applied the Bernoulli-sampling method to the

DQPS protocol, which was recently proved to be secure in the asymptotic regime. Although the asymptotic proof is based on the tagging of the insecure rounds as in the WCP-BB84 protocol, the definition of the tagged round is much more convoluted and makes sense only after the signal was detected by Bob. Nonetheless, our finite-key analysis has led to a key rate formula closely analogous to the one for the WCP-BB84 protocol. Numerical calculation shows that the DQPS protocol retains higher key rates than the BB84 protocol with phase encoding (PE-BB84) even in the finite-key regime of  $n_{\text{rep}} = 10^7$ .

It is expected that our method can also be applied to protocols with decoy states [27–29]. Since the existing analyses [6–9] with decoy states involve the estimation of the sample size  $n_{X,\text{unt}}$ , the present method may provide a simpler analysis compared to the conventional methods with simple random sampling. It should be mentioned that some of the finite key analyses [6, 7] assumed the announcement of basis choice after each round to make the sample size fixed, which were later pointed out [30] to open a security hole against a sifting attack. This illustrates an importance of simpler and more straightforward methods, and we believe that the method proposed here will contribute in this regard.

## ACKNOWLEDGEMENT

We thank A. Mizutani, T. Tsurumaru and K. Yoshino for helpful discussions. This work was supported by the ImPACT Program of the Council for Science, Technology and Innovation (Cabinet Office, Government of Japan), CREST, and the Photon Frontier Network Program (MEXT).

## Appendix: Description of the DQPS protocol

Here we summarize the detail of the DQPS protocol in Ref. [13]. The protocol proceeds as follows, which includes predetermined parameters  $\tilde{p}_X > 0$ ,  $\tilde{p}_Z = 1 - \tilde{p}_X$ ,  $L \geq 2$ , and  $n_{\text{rep}}$ .

1. Alice selects a bit  $c \in \{0, 1\}$  with probability  $\tilde{p}_Z$  and  $\tilde{p}_X$ , which correspond to the choice of  $Z$  basis and  $X$  basis, respectively. Bob also selects  $d \in \{0, 1\}$  with probability  $\tilde{p}_Z$  and  $\tilde{p}_X$ .
2. Alice generates  $L$  random bits  $a_l \in \{0, 1\}$  ( $l = 0, 1, \dots, L-1$ ), and prepares  $L$  optical pulses (system  $S$ ) in the state

$$\hat{\rho}_S = \hat{S}(\{a_l\}, c) \hat{\sigma}_S \hat{S}(\{a_l\}, c) \quad (A.1)$$

$$\hat{S}(\{a_l\}, c) := \bigotimes_{l=0}^{L-1} \exp(i(a_l \pi + \frac{\pi}{2} l c) \hat{n}_l),$$

where  $\hat{\sigma}_S$  is the state of the  $L$  pulses from the source before phase modulation and  $\hat{n}_l$  represents the photon number operator for the  $l$ -th pulse. Alice randomizes the overall optical phase of the  $L$ -pulse train, and sends it to Bob.

3. If  $d = 0$ , Bob sets the phase shift  $\theta_B = 0$ . If  $d = 1$ , he sets  $\theta_B = \frac{\pi}{2}$ .
4. If there is no detection of photons at the valid timings, Bob



sets  $j = 0$ . If the detections have only occurred at a single valid timing, the variable  $j$  is set to the index of the timing. If there are detections at multiple timings, the smallest (earliest) index of them is assigned to  $j$ . If  $j \neq 0$ , Bob determines his raw key bit  $b \in \{0, 1\}$  depending on which detector has reported detection at the  $j$ -th timing. If both detectors have reported at the  $j$ -th timing, a random bit is assigned to  $b$ . Bob announces  $j$  publicly.

5. If  $j \neq 0$ , Alice determines her raw key bit as  $a = a_{j-1} \oplus a_j$ .

6. Alice and Bob repeat the above procedures  $n_{\text{rep}}$  times. They publicly disclose  $c$  and  $d$  for each of the  $n_{\text{rep}}$  rounds.

7. Alice and Bob define bit strings  $\kappa_{A,X}$  and  $\kappa_{B,X}$ , respectively, by concatenating their determined bits with  $j \neq 0$  and  $c = d = 1$ . They define sifted keys  $\kappa_{A,Z}$  and  $\kappa_{B,Z}$ , respectively, by concatenating their determined bits with  $j \neq 0$  and  $c = d = 0$ . Let their sizes be  $n_Z := |\kappa_{A,Z}| = |\kappa_{B,Z}|$  and  $n_X := |\kappa_{A,X}| = |\kappa_{B,X}|$ .

8. They disclose and compare  $\kappa_{A,X}$  and  $\kappa_{B,X}$  to determine the number of bit errors  $k_X$  included in them.

9. Through public discussion, Bob corrects his keys  $\kappa_{B,Z}$  to make it coincide with Alice's key  $\kappa_{A,Z}$  and obtains  $\kappa_{B,Z}^{\text{cor}}$  ( $|\kappa_{B,Z}^{\text{cor}}| = n_Z$ ).

10. Alice and Bob conduct privacy amplification by shortening  $\kappa_{A,Z}$  and  $\kappa_{B,Z}^{\text{cor}}$  to obtain final keys  $\kappa_{A,Z}^{\text{fin}}$  and  $\kappa_{B,Z}^{\text{fin}}$  of size  $n_{\text{fin}}$ .

The security of the above protocol in the asymptotic limit was proved in [13] under the following assumptions on the devices used by Alice and Bob. The discussion in subsection IV C of the main text uses the same assumptions. We assume that the phase randomization in Step 2 is ideal, and hence the state emitted from Alice in Step 2 is expressed as

$$\sum_m \hat{N}_m \hat{\rho}_S \hat{N}_m, \quad (\text{A.2})$$

where  $\hat{N}_m$  represents the projector onto the subspace with  $m$  total photons in the  $L$  pulses. We also assume that a parameter

$r_{\text{tag}}$  associated with the  $L$ -pulse state  $\hat{\rho}_S$  from the source is known or at least is bounded from above. With  $|m_l\rangle_{S,l}$  being an  $m$ -photon state of the  $l$ -th pulse, the parameter is defined by Eq. (85) with

$$\hat{\Pi}_S^{(m)} := \sum_{\{m_l\} \in \Gamma^{(m)}} \bigotimes_{l=0}^{L-1} |m_l\rangle \langle m_l|_{S,l}, \quad (\text{A.3})$$

where  $\Gamma^{(m)}$  is a set of values of  $L$  nonnegative integers

$$\Gamma^{(m)} := \left\{ (i_0, \dots, i_{L-1}) \mid i_{l-1} + i_l \leq 1 (1 \leq l \leq L-1), \sum_{l=0}^{L-1} i_l = m \right\}. \quad (\text{A.4})$$

In [13], we showed a practical method of off-line calibration to determine an upper bound of  $r_{\text{tag}}$  for a general light source.

To describe the assumptions for Bob's apparatus, we introduce POVM elements for Bob's procedure in Steps 3 and 4. Let  $\{\hat{B}_j^{(d)}\}_{j=0, \dots, L-1}$  be the POVM for Bob's procedure of determining  $j$ , when the basis  $d$  was selected in Step 1. We further decompose the elements for  $j \neq 0$  as  $\hat{B}_j^{(d)} = \hat{B}_{j,0}^{(d)} + \hat{B}_{j,1}^{(d)}$ , where  $\hat{B}_{j,b}^{(d)}$  corresponds to the outcome  $(j, b)$ . These operators satisfy

$$\hat{B}_0^{(d)} + \sum_{j=1}^{L-1} (\hat{B}_{j,0}^{(d)} + \hat{B}_{j,1}^{(d)}) = \hat{\mathbb{I}}. \quad (\text{A.5})$$

We then assume that Bob uses threshold detectors, and further assume that the inefficiency and dark countings of the detectors are equivalently represented by an absorber and a stray photon source placed in front of Bob's apparatus, and hence they are included in the quantum channel. This leads [13] to the condition

$$\hat{B}_j^{(0)} = \hat{B}_j^{(1)} \quad (0 \leq j \leq L-1). \quad (\text{A.6})$$

- 
- [1] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, Vol. 175, Bangalore, India (IEEE Press, New York, 1984).
  - [2] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, in *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, Lecture Notes in Computer Science, Vol. 3378 (Springer, 2005) pp. 386–406.
  - [3] R. Renner and R. König, in *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, Lecture Notes in Computer Science, Vol. 3378 (Springer, 2005) pp. 407–425.
  - [4] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, *Nature Communications* **3**, 634 (2012).
  - [5] F. Furrer, T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner, *Phys. Rev. Lett.* **109**, 100502 (2012).
  - [6] C. C. W. Lim, M. Curty, N. Walenta, F. Xu, and H. Zbinden, *Phys. Rev. A* **89**, 022307 (2014).
  - [7] M. Curty, F. Xu, W. Cui, C. C. W. Lim, K. Tamaki, and H.-K. Lo, *Nature Communications* **5**, 3732 (2014).
  - [8] M. Hayashi and R. Nakayama, *New Journal of Physics* **16**, 063009 (2014).
  - [9] M. Lucamarini, J. F. Dynes, B. Fröhlich, Z. Yuan, and A. J. Shields, *IEEE Journal of Selected Topics in Quantum Electronics* **21**, 197 (2015).
  - [10] A. Mizutani, M. Curty, C. C. W. Lim, N. Imoto, and K. Tamaki, *New Journal of Physics* **17**, 093011 (2015).
  - [11] H.-K. Lo, H. Chau, and M. Ardehali, *Journal of Cryptology* **18**(2), 133 (2004).
  - [12] K. Inoue and Y. Iwai, *Phys. Rev. A* **79**, 022319 (2009).
  - [13] S. Kawakami, T. Sasaki, and M. Koashi, *Phys. Rev. A* **94**, 022332 (2016).
  - [14] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, Vol. 175, Bangalore, India (IEEE Press, New York, 1984).
  - [15] C. H. Bennett, G. Brassard, and N. D. Mermin, *Phys. Rev. Lett.* **68**, 557 (1992).



- [16] H.-K. Lo and H. F. Chau, *Science* **283**, 2050 (1999).
- [17] P. W. Shor and J. Preskill, *Phys. Rev. Lett.* **85**, 441 (2000).
- [18] W. Hoeffding, *Journal of the American Statistical Association* **58**, 13 (1963).
- [19] J. H. Ahrens, “A comparison of hypergeometric distributions with corresponding binomial distributions,” in *Ökonomie und Mathematik: Rudolf Henn zum 65. Geburtstag*, edited by O. Opitz and B. Rauhut (Springer Berlin Heidelberg, Berlin, Heidelberg, 1987) pp. 253–265.
- [20] D. Gottesman, H.-K. Lo, J. Preskill, and N. Lütkenhaus, *Quant. Info. Compu.* **5**, 325 (2004).
- [21] T. Tsurumaru and K. Tamaki, *Phys. Rev. A* **78**, 032302 (2008).
- [22] N. J. Beaudry, T. Moroder, and N. Lütkenhaus, *Phys. Rev. Lett.* **101**, 093601 (2008).
- [23] M. Koashi, *New J. Phy.* **11**, 045018 (2009).
- [24] R. Y. Q. Cai and V. Scarani, *New Journal of Physics* **11**, 045024 (2009).
- [25] H. Chernoff, *Ann. Math. Stat.* **23**, 493 (1952).
- [26] The argument in Ref. [13] represented various events in  $n_{\text{rep}}$  rounds through strings  $\mathbf{c}, \mathbf{a}, \mathbf{b}, \mathbf{j}$  and  $\mathbf{t}$  of length  $n_{\text{rep}}$ , and proved that their joint probability distribution takes a form  $\Pr(\mathbf{c}, \mathbf{a}, \mathbf{b}, \mathbf{j}, \mathbf{t}) = \Pr(\mathbf{c})\beta(g_{t,j}(\mathbf{c}), \mathbf{a}, \mathbf{b}, \mathbf{j}, \mathbf{t})$  as stated after Eq.(A3) of Ref. [13]. In rewriting it to Eq. (72) of the main text, we used the following facts. There is a one-to-one correspondence between  $\mathcal{X}_A$  and  $\mathbf{c}$ .  $\mathcal{N}_{\text{unt}}$  and  $\mathcal{K}_{\text{unt}}$  are functions of  $\mathbf{a}, \mathbf{b}, \mathbf{j}$  and  $\mathbf{t}$ .  $g_{t,j}(\mathbf{c})$  is a function of  $\{\mathcal{X}_A \cap \overline{\mathcal{N}_{\text{unt}}}, \mathcal{N}_{\text{unt}}\}$ .
- [27] W.-Y. Hwang, *Phys. Rev. Lett.* **91**, 057901 (2003).
- [28] X.-B. Wang, *Phys. Rev. Lett.* **94**, 230503 (2005).
- [29] H.-K. Lo, X. Ma, and K. Chen, *Phys. Rev. Lett.* **94**, 230504 (2005).
- [30] C. Pfister, N. Lutkenhaus, S. Wehner, and P. J. Coles, *New Journal of Physics* **18**, 053001 (2016).