Security of decoy-state protocols for general photon-number-splitting attacks

Rolando D. Somma* and Richard J. Hughes†
Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
(Received 22 April 2013; published 24 June 2013)

Decoy-state protocols provide a way to defeat photon-number-splitting attacks in quantum cryptography implemented with weak coherent pulses. We point out that previous security analyses of such protocols relied on assumptions about eavesdropping attacks that considered treating each pulse equally and independently. We give an example to demonstrate that, without such assumptions, the security parameters of previous decoy-state implementations could be worse than the ones claimed. Next we consider more general photon-number-splitting attacks, which correlate different pulses, and give an estimation procedure for the number of single-photon signals with rigorous security statements. The impact of our result is that previous analyses of the number of times a decoy-state quantum cryptographic system can be reused before it makes a weak key must be revised.

DOI: 10.1103/PhysRevA.87.062330 PACS number(s): 03.67.Dd, 03.67.Hk, 03.67.Ac, 42.50.−p

I. INTRODUCTION

Quantum key distribution (QKD) [1–4] allows two parties, Alice and Bob, to establish a common and secret key S that is informationally secure; see [5–7] and references therein. A widely used setup for QKD is the one suggested by Bennett and Brassard (BB84) [2]. BB84 is ideally implemented by preparing and transmitting single-photon pulses. Information can be encoded in the state of one of two conjugate polarization bases, e.g., vertical/horizontal or diagonal/antidiagonal. Only those photons that were prepared by Alice and detected by Bob in the same basis are useful to build a sifted key, which forms S after additional steps of information reconciliation and privacy amplification. Security follows from the inability of faithfully copying quantum information [8] and the unavoidable information-disturbance trade-off in quantum mechanics. Nevertheless, realistic implementations of BB84 use weak coherent photon pulses that could involve many photons, avoiding the assumptions made in security analyses [9–11]. Such pulses could be exploited by Eve, the eavesdropper, to gain access to the (insecure) distributed key using a so-called photon-number-splitting (PNS) attack [12,13]. In a simple proposed PNS attack, Eve measures the number of photons in the pulse, n. If $n = 1$, Eve blocks the pulse. If $n \geqslant 2$, Eve "splits" the pulse to obtain a copy of a single photon with the correct polarization and keeps it in her quantum memory. Eve could then obtain a full copy of S by making measurements of her photons in the correct polarization bases, which are known after a public discussion between Alice and Bob. Since Alice and Bob cannot measure n, a PNS attack may go undetected. Our goal is to provide a protocol for secure QKD in the presence of PNS attacks.

A simple approach to overcome a PNS attack considers reducing the probability of multiphoton pulses by reducing the coherent-pulse intensities. The drawback with this approach is that the probability of creating single-photon pulses is also reduced. Then, the rate at which bits to build S are sifted is far from optimal [13,14]. Another approach is to use decoy states that allow one to detect PNS attacks without a substantial reduction of the rate of sifted bits if Eve is not present [15–17]. In a decoy-state protocol (DSP), one of several weak coherent sources is randomly selected for each pulse. Such sources create pulses of different intensities (mean photon numbers). This gives Alice and Bob a means to estimate f_0 and f_1, the number of Bob's detections due to empty and single-photon pulses prepared by Alice, in the same basis, respectively. The values of f_0 and f_1 are important to determine $|S|$, the length of the secure key. For example, in the discussed PNS attack, f_1 is substantially smaller than its value when Eve is not present, and so is $|S|$.

In more detail, we let $K \gg 1$ be the total number of pulses prepared by Alice. We first assume that the channel is nonadversarial, i.e., no eavesdropping attacks are present. If the pulse has a random phase, the number of photons it contains is sampled according to the Poisson distribution:

$$p_n^\mu = \Pr(n|\mu) = e^{-\mu}\frac{\mu^n}{n!}, \tag{1}$$

where μ is the mean photon number for that source and $\mu \leqslant 1$ in applications. We let η be the transmission/detection efficiency of the quantum channel shared by Alice and Bob. If $b = 1$ ($b = 0$) denotes the event in which Bob detects a nonempty (empty or vacuum) pulse,

$$y_n = \Pr(b = 1|n)$$

is the probability of a detection by Bob given that Alice's prepared pulse contained n photons. y_n is the so-called n-photon yield and $y_n < 1$ due to losses in the channel. For $n \geqslant 1$, we may assume

$$y_n = 1 - (1 - \eta)^n,$$

which is a good approximation in applications. For $n = 0$, $y_0 > 0$ denotes Bob's detector dark-count rate. The total probability of Bob detecting a pulse (in any one cycle) is the total yield

$$Y(\mu) = \Pr(b = 1) = \sum_{n\geqslant 0} \Pr(n|\mu) y_n = e^{-\mu} y_0 + 1 - e^{-\mu\eta}. \tag{2}$$

$Y(\mu)$ can be estimated by Alice and Bob, via public discussion, from the frequency of detections after all pulses were transmitted.

*somma@lanl.gov
†rxh@lanl.gov

In QKD, we allow Eve to manipulate the parameters that characterize the channel at her will. We use the superscript $\mathcal{E}$ to represent the interaction of Eve with the communication. For example, $y_n^{\mathcal{E}}$ denotes the n-photon yield in the presence of Eve. In a general intercept-resend attack, Eve may intercept a pulse and resend a different one. That is, each detection by Bob is not guaranteed to come from the same pulse that Alice prepared. In a simple PNS attack, Eve makes nondemolition measurements of n. With this information, Eve sets $y_1^{\mathcal{E}} = 0 \neq y_1$ and $y_n^{\mathcal{E}} \geqslant y_n$ for $n \geqslant 2$, so that

$$Y(\mu) \approx Y^{\mathcal{E}}(\mu).$$

Then, if Alice and Bob can only estimate the total yields, a PNS attack could be "invisible" with the right choices of $y_2^{\mathcal{E}}, y_3^{\mathcal{E}}, \ldots$. To increase the multiphoton yield, Eve may use an ideal channel to resend the pulses. (Note that sophisticated PNS attacks that do not change the Poisson distribution are possible [13].) A PNS attack allows Eve to have the full key S if

$$\Pr(n \geqslant 2|\mu) \geqslant Y(\mu).$$

In this case, Eve possesses a photon with the same polarization as that of the pulse detected by Bob and no single-photon pulses are involved in creating S. Only if $\Pr(n \geqslant 2|\mu) < Y(\mu)$ are some security guarantees possible [14]. Such an inequality is satisfied when $\mu \approx \eta$, implying a rate for sifted bits of order η^2 [Eq. (2)]. This is undesirably small ($\eta \ll 1$).

Remarkably, DSPs give an optimal rate of order η with small resource overheads. A goal in a DSP is to estimate $y_0^{\mathcal{E}}$ and $y_1^{\mathcal{E}}$, which provide a lower bound on $f_0^{\mathcal{E}}$ and $f_1^{\mathcal{E}}$, respectively. Empty and single-photon pulses cannot be split and the information carried in their polarization cannot be faithfully copied, making them useful for creating a secure key. For the estimation, Alice uses photon sources with different values of μ, but they are identical otherwise. In a conventional DSP, it is assumed that Eve's PNS attack treats every n-photon pulse equally and independently, regardless of its source. That is, Eve's attack is simulated by independent and identically distributed (i.i.d.) random variables. The total yield in this case is, for any given μ,

$$Y^{\mathcal{E}}(\mu) = \sum_{n \geqslant 0} p_n^{\mu}\, y_n^{\mathcal{E}}. \tag{3}$$

Equation (3) describes mathematically what we denote as the i.i.d. assumption. It follows that

$$y_0^{\mathcal{E}} = Y^{\mathcal{E}}(\mu)|_{\mu=0}, \quad y_1^{\mathcal{E}} = \partial_\mu [e^{\mu} Y^{\mathcal{E}}(\mu)]|_{\mu=0}. \tag{4}$$

Then, if Eve's attack satisfies $Y(\mu) \approx Y^{\mathcal{E}}(\mu)$ for all μ,

$$y_0^{\mathcal{E}} \approx y_0, \quad y_1^{\mathcal{E}} \approx y_1 = \eta.$$

That is, by being able to estimate $Y^{\mathcal{E}}(\mu)$ for two values of $\mu \ll 1$ via public discussion, Alice and Bob can restrict Eve's attack so that the dark-count rate and single-photon yield are almost unchanged from the nonadversarial case. In addition, if a third source with $\mu \approx 1$ is randomly invoked, an optimal key rate of order η will be achieved.

In reality, the estimation of $y_0^{\mathcal{E}}$ and $y_1^{\mathcal{E}}$ is subject to finite statistics and can be technically involved. Nevertheless, the i.i.d. assumption in Eq. (3) allows Alice and Bob to gain information about Eve's attack by running the protocol and analyzing the (binomial) distributions of the detection events for each source. However, we remark that if Eve were to correlate her attacks, the i.i.d. assumption and the corresponding security analyses would be invalid. This is the main motivation behind our analysis.

In this paper, we give an example that shows how the i.i.d. assumption can be simply bypassed by Eve, resulting in security parameters that are worse from those obtained under the assumption. We then analyze the security of DSPs for general PNS attacks. Our main result is an estimation procedure that gives a lower bound on $f_0^{\mathcal{E}}$ and $f_1^{\mathcal{E}}$, with a confidence level that is an input to the estimation procedure. Our security analysis does not use the i.i.d. assumption and is particularly relevant when Eve performs a PNS attack that could correlate different pulses in one session or even different sessions. We compare some results obtained by our estimation procedure with those obtained by using the i.i.d. assumption, and we emphasize the importance of our procedure.

II. THE SECURITY PARAMETER, THE I.I.D. ASSUMPTION, AND FINITE STATISTICS

Of high significance in cryptographic protocols is ϵ, the so-called security parameter. ϵ measures the deviation of a real protocol implementation from an ideal one. We use the same definition used in Ref. [7], which states that a real QKD protocol is ϵ-secure if it is ϵ-indistinguishable from a perfectly secure and ideal one. This definition is equivalent to a statement on the trace norm of the difference between the quantum states resulting from the real and ideal protocols, respectively. It implies that a QKD protocol that is ϵ-secure can be safely reused order $1/\epsilon$ times without compromising its security.

Usually, one fixes a value for ϵ and then determines the size of S based on several protocol performance parameters. These parameters include the number of pulses sent by Alice, the number of pulses detected by Bob, and the estimated bit error rates at each mean photon number. For DSPs, ϵ has a component ϵ_{DSP} that determines the confidence level in the estimation of a lower bound of $f_0^{\mathcal{E}}$ and $f_1^{\mathcal{E}}$, due to finite statistics.

A possible way to obtain such lower bounds, under the i.i.d. assumption, is the one followed in Ref. [17]. In this case, we consider a DSP with three sources, $i = U, V, W$. The mean photon number in each pulse, for each source, is $\mu^U = 0$, $\mu^V \ll 1$, and $\mu^W \in O(1)$. Each source i randomly prepares a pulse with probability q^i and we let K^i be the total number of pulses for that source. K^i is known to Alice and Bob by public discussion after all pulses are sent and $K^i \approx q_i K$ when $K \gg 1$. We write $D^{i,\mathcal{E}}$ for the random variable that counts the number of pulses from source i detected by Bob under the presence of Eve [18]. The exact value that $D^{i,\mathcal{E}}$ takes in a session can also be obtained by Alice and Bob via public discussion after the pulses were transmitted.

Under the i.i.d. assumption [Eq. (3)], $D^{i,\mathcal{E}}$ is sampled according to the binomial distribution. Then, $D^{i,\mathcal{E}}/K^i$ is an estimator of the total yield $Y^{\mathcal{E}}(\mu^i) = E[D^{i,\mathcal{E}}/K^i]$, where $E[.]$ denotes the mean value. That is, for a given $\bar{\epsilon}_{\mathrm{DSP}}$, we can establish confidence intervals

$$\frac{D^{i,\mathcal{E}}}{K^i} + c(\bar{\epsilon}_{\mathrm{DSP}})\sigma^{i,\mathcal{E}} \geqslant Y^{\mathcal{E}}(\mu^i) \geqslant \frac{D^{i,\mathcal{E}}}{K^i} - c(\bar{\epsilon}_{\mathrm{DSP}})\sigma^{i,\mathcal{E}}, \tag{5}$$

with confidence level $1-\bar{\epsilon}_{\mathrm{DSP}}$. The constant c depends on $\bar{\epsilon}_{\mathrm{DSP}}$ and can be obtained by using Chernoff's bound [19] (see Appendix B). The standard deviation in this case is

$$\sigma^{i,\mathcal{E}} \approx \sqrt{\frac{Y^{\mathcal{E}}(\mu^i)[1-Y^{\mathcal{E}}(\mu^i)]}{q^i K}}. \tag{6}$$

Using Eq. (3) for $Y^{\mathcal{E}}(\mu^i)$, we can search for the minimum values of $y_0^{\mathcal{E}}$ and $y_1^{\mathcal{E}}$ that satisfy Eqs. (5), e.g., by executing a linear program. Both $y_0^{\mathcal{E}}$ and $y_1^{\mathcal{E}}$ can then be used to obtain the desired lower bounds on $f_0^{\mathcal{E}}$ and $f_1^{\mathcal{E}}$, respectively, with corresponding confidence level $1-\epsilon_{\mathrm{DSP}}$. This last step also requires using the i.i.d. assumption [20].

III. INCREASING THE LENGTH OF CONFIDENCE INTERVALS: AN ATTACK

The analysis in Sec. II used the i.i.d. assumption that resulted in a value for $\sigma^{i,\mathcal{E}}$ given by Eq. (6). Nevertheless, the actual value of $\sigma^{i,\mathcal{E}}$ could be much higher in more general PNS attacks. For the same confidence level, a bigger $\sigma^{i,\mathcal{E}}$ implies a "wider" confidence interval for the estimation of the yield $Y^{\mathcal{E}}(\mu^i)$ (Appendix B) and thus smaller lower bounds on $f_0^{\mathcal{E}}$ and $f_1^{\mathcal{E}}$. The overall result in the DSP is a secret key S of smaller size for the same security parameter.

To illustrate how Eve can bypass the i.i.d. assumption, we suggest a potential attack that results in almost no change for the total yields [i.e., $Y(\mu^i) \approx Y^{\mathcal{E}}(\mu^i)$] [21] but the variances $\sigma^{i,\mathcal{E}}$ are increased with respect to those of the binomial distribution [Eq. (6)]. The suggested attack could be detected by Alice and Bob by estimating the variances directly via public discussion. Nevertheless, it still shows that a better analysis of the security of DSPs is needed to make rigorous claims.

In the attack, Eve first picks an integer value for $\tau \geqslant 1$, where τ^2 denotes a scale for a variance or "correlation" of a particular distribution. Eve receives all pulses from Alice and we let k_n be the total number of n-photon pulses in the protocol. Note that the exact value of k_n is known to Eve but not to Alice and Bob. In general, k_n is sampled according to the binomial distribution

$$\Pr(k_n) = \binom{K}{k_n} (p_n)^{k_n} (1-p_n)^{K-k_n},$$

where p_n is the probability of a pulse containing n photons: $p_n = \sum_i q^i p_n^{\mu^i}$. The mean and variance for such distribution are

$$E[k_n] = p_n K, \quad \sigma_{k_n}^2 = p_n(1-p_n)K.$$

Given k_n, Eve randomly picks a value for $d_n^{\mathcal{E}} \in \{0,1,\ldots,k_n\}$, where $d_n^{\mathcal{E}} = \sum_i d_n^{i,\mathcal{E}}$ is the total number of detections due to n-photon pulses prepared by Alice. In particular, we assume that Eve can control $d_0^{\mathcal{E}}$, which determines the dark-count rate. The distribution associated with $d_n^{\mathcal{E}}$ has the following properties:

$$E[d_n^{\mathcal{E}}|k_n] = y_n k_n, \quad \sigma^2_{d_n^{\mathcal{E}}|k_n} = \tau^2 y_n(1-y_n)k_n. \tag{7}$$

We let $d_n^{i,\mathcal{E}}$ be the number of n-photon pulses, prepared by Alice's ith source only and detected by Bob. The exact value of $d_n^{i,\mathcal{E}}$ is unknown to all parties. Because Eve does not know the source being used in the DSP, $d_n^{i,\mathcal{E}}$ is sampled according to the binomial distribution when given $d_n^{\mathcal{E}}$:

$$\Pr\left(d_n^{i,\mathcal{E}}|d_n^{\mathcal{E}}\right) = \binom{d_n^{\mathcal{E}}}{d_n^{i,\mathcal{E}}} \left(q_n^i\right)^{d_n^{i,\mathcal{E}}} \left(1-q_n^i\right)^{d_n^{\mathcal{E}}-d_n^{i,\mathcal{E}}}, \tag{8}$$

where

$$q_n^i = \frac{q^i e^{-\mu^i}(\mu^i)^n}{\sum_{i'=U,V,W} q^{i'} e^{-\mu^{i'}}(\mu^{i'})^n}. \tag{9}$$

The distribution associated with $d_n^{i,\mathcal{E}}$ satisfies

$$E\left[d_n^{i,\mathcal{E}}|d_n^{\mathcal{E}}\right] = q_n^i d_n^{\mathcal{E}}, \quad \sigma^2_{d_n^{i,\mathcal{E}}|d_n^{\mathcal{E}}} = q_n^i\left(1-q_n^i\right)q_n^{\mathcal{E}}.$$

As in Sec. II, we let $(\sigma^{i,\mathcal{E}})^2$ be the variance associated with the random variable $Z^{i,\mathcal{E}} = D^{i,\mathcal{E}}/K^i$, where

$$D^{i,\mathcal{E}} = \sum_{n\geqslant 0} d_n^{i,\mathcal{E}} \tag{10}$$

and $E[Z^{i,\mathcal{E}}] = Y^{\mathcal{E}}(\mu^i)$. An accurate estimate of $Z^{i,\mathcal{E}}$ can be obtained if we approximate $K^i \approx q^i K$, in the limit of large K. In addition, because K is fixed, the variables k_n are not independent. However, in the large-K limit, k_n can also be approximated by its mean value. This implies that the k_n are almost independent and so are the $d_n^{\mathcal{E}}$ and $d_n^{i,\mathcal{E}}$ for different values of n. Under these approximations,

$$(\sigma^{i,\mathcal{E}})^2 \approx \frac{1}{(q^i K)^2} \sum_{n\geqslant 0} \sigma^2_{d_n^{i,\mathcal{E}}}. \tag{11}$$

In Appendix A we show that

$$\sigma^2_{d_n^{i,\mathcal{E}}} = \left[(\tau^2-1)q_n^i(1-y_n) + \left(1-q_n^i y_n p_n\right)\right] q_n^i y_n p_n K. \tag{12}$$

By inserting Eq. (12) in Eq. (11), we can obtain the variances as a function of τ. In Fig. 1 we compute $\sigma^{U,\mathcal{E}}$ and $\sigma^{V,\mathcal{E}}$. The i.i.d. assumption discussed in Sec. II corresponds to $\tau = 1$ (see Appendix A). Using these results in Eq. (5) yields wider confidence intervals for the same confidence level.

To illustrate our point further, we consider a simple protocol in which a single source U is used to estimate the dark-count rate. Here, $\mu^U = 0$ and $d_0^{\mathcal{E}} = D^{U,\mathcal{E}}$ is known.

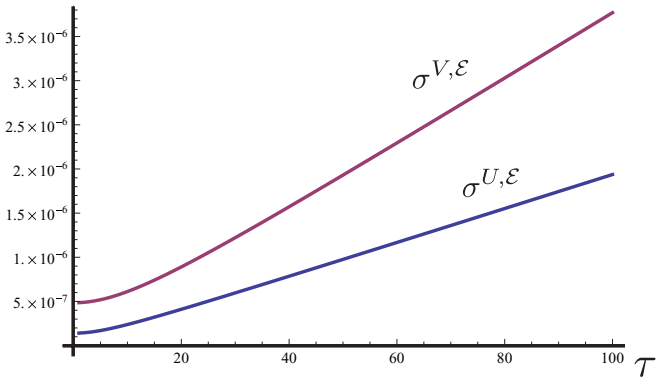

FIG. 1. (Color online) The standard deviations $\sigma^{U,\mathcal{E}}$ and $\sigma^{V,\mathcal{E}}$ for an attack in which Eve correlates n-photon pulses according to the value of τ. The channel parameters are $1 \leqslant \tau \leqslant 100$, $K = 10^{10}$, $q^U = 0.01$, $q^V = 0.0275$, $\mu^U = 0$, $\mu^V = 0.063$, $\eta = 10^{-3}$, and $y_0 = 2\times 10^{-6}$ [17]. The results in Sec. II are recovered for $\tau = 1$.

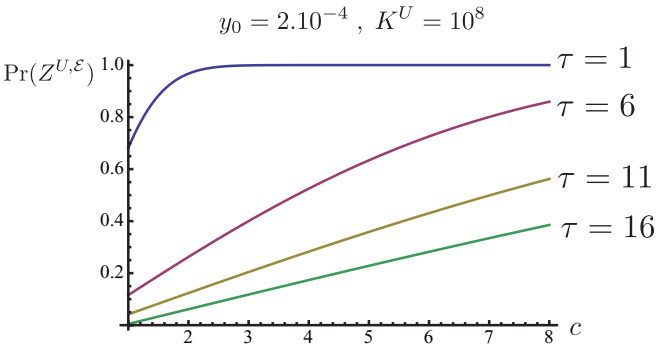

FIG. 2. (Color online) Estimation of dark counts: confidence intervals for different correlated attacks, parametrized by τ, and confidence bounds, parametrized by c.

In the nonadversarial case, $d_0^{\mathcal{E}}$ is sampled according to the binomial distribution with probability y_0 and known sample size $k_0 = K^U$. Nevertheless, for the correlated attack, we assume that Eve "receives" the K^U pulses and groups them according to blocks of size τ^2. Then, Eve will force (prevent) the detection of all pulses in any one block with probability $y_0\,(1-y_0)$. The random variable $d_0^{\mathcal{E}}$ for the correlated attack satisfies

$$E[d_0^{\mathcal{E}}] = y_0 k_0,$$

$$\sigma^2_{d_0^{\mathcal{E}}}(\tau) = [y_0(\tau^2)^2 - (y_0\tau^2)^2]\frac{k_0}{\tau^2} = \tau^2 y_0(1-y_0)k_0,$$

and $\tau = 1$ corresponds again to the i.i.d. assumption [see Eq. (7)]. In Fig. 2, we plot the probability that $Z^{U,\mathcal{E}} = D^{U,\mathcal{E}}/K^U$ satisfies

$$E[Z^{U,\mathcal{E}}] + c\sigma_{d_0^{\mathcal{E}}}(1) \geqslant Z^{U,\mathcal{E}} \geqslant E[Z^{U,\mathcal{E}}] - c\sigma_{d_0^{\mathcal{E}}}(1),$$

for different values of c and τ. For $\tau = 1$, such a probability corresponds to the confidence level in Eq. (5). $E[Z^{U,\mathcal{E}}] = y_0$ in this example.

Our results demonstrate that, for a fixed security parameter, the accuracy in the estimation of the dark-count rate strongly depends on Eve's attack and can be substantially different from the one obtained under the i.i.d. assumption ($\tau = 1$).

IV. SECURITY OF DSP: CORRELATED PNS ATTACKS

We go beyond the i.i.d. assumption and study more general and correlated PNS attacks, in which Eve has full control of Bob's detection events. The secure key rate in a realistic implementation of QKD is [17]

$$s \geqslant f_0^{\mathcal{E}*} + f_1^{\mathcal{E}*} - \kappa_{\mathrm{EC}} F^{\mathcal{E}} H_2(\mathrm{BER}) - \kappa_{\mathrm{PA}} f_1^{\mathcal{E}*} H_2\big(b_1^{\max}\big), \quad (13)$$

which determines the size of the distributed key as $|S| = sK$. $F^{\mathcal{E}}$ is the total number of pulses detected by Bob and prepared by Alice in the same basis, which are useful for the sifted key. In BB84, $F^{\mathcal{E}} \approx D^{\mathcal{E}}/2$, where $D^{\mathcal{E}}$ is the total number of detections. $f_n^{\mathcal{E}*}$ is a lower bound on $f_n^{\mathcal{E}}$, the number of n-photon pulses prepared and detected in the same basis. $H_2(.)$ is the Shannon entropy, κ_{EC} and κ_{PA} are coefficients due to the error correction and privacy amplification steps, BER is the total bit error rate, and $b_1^{\max}$ is an upper bound to the bit error rate due to single-photon pulses only.

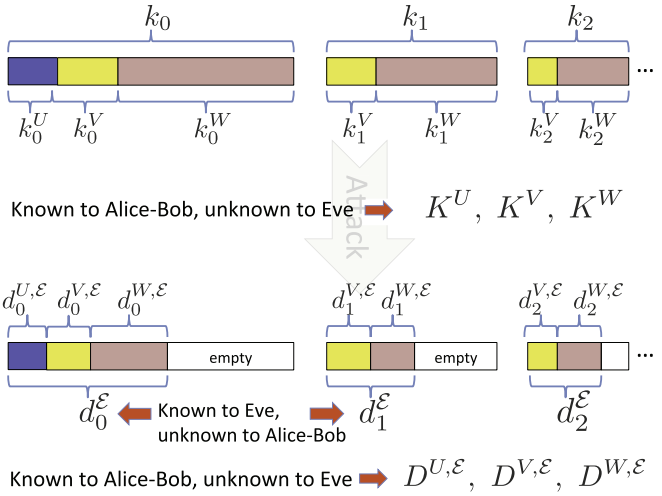

FIG. 3. (Color online) A general PNS attack with three decoy sources: $\mu_U = 0$, $\mu_V \ll 1$, and $\mu_W \in O(1)$. Each block represents the number of pulses with $n = 0,1,2,\ldots$, respectively. The random variables k_n indicate the number of n-photon pulses prepared by Alice and the superscript i denotes the source used for such pulses. Eve's attack controls the number of detections by Bob, due to n-photon pulses, through $d_n^{\mathcal{E}}$.

In a DSP, we characterize a general PNS attack by the distribution

$$\Pr\big(d_0^{\mathcal{E}}, d_1^{\mathcal{E}}, \ldots | k_0, k_1, \ldots\big); \quad (14)$$

see Fig. 3 for an example. Our goal is to build an estimation procedure that places confidence intervals on $f_0^{\mathcal{E}} = \sum_i f_0^{i,\mathcal{E}}$ and $f_1^{\mathcal{E}} = \sum_i f_1^{i,\mathcal{E}}$ from the known $D^{i,\mathcal{E}}$. These intervals ultimately imply a lower bound on s [see Eq. (13)].

We assume that there are three sources satisfying $\mu_U = 0 < \mu_V < \mu_W$, and $\mu_W \in O(1)$. Nevertheless, our analysis can be easily generalized to the case in which more sources are present, where the estimation is more accurate. For each source, Bob's detections satisfy Eq. (10). If a simple relationship between each $d_n^{i,\mathcal{E}}$ and $d_n^{\mathcal{E}}$ could be found, we could execute a program to solve Eqs. (10). Such a relationship could be obtained from the binomial distribution associated with $d_n^{i,\mathcal{E}}$, when given $d_n^{\mathcal{E}}$ [Eq. (8)] (see below).

Our estimation procedure uses $d_n^{i,\mathcal{E}}$ to determine the confidence intervals

$$\Phi_{i,n}\big(d_n^{i,\mathcal{E}}\big) \geqslant d_n^{\mathcal{E}} \geqslant \phi_{i,n}\big(d_n^{i,\mathcal{E}}\big). \quad (15)$$

The corresponding confidence level for each inequality is $1 - \epsilon_n/2$. The upper and lower bounds are monotonic and invertible functions. Then,

$$\phi_{i,n}^{-1}\big(d_n^{\mathcal{E}}\big) \geqslant d_n^{i,\mathcal{E}} \geqslant \Phi_{i,n}^{-1}\big(d_n^{\mathcal{E}}\big), \quad (16)$$

with the same confidence levels. Such confidence levels do not result from the binomial distribution as we are analyzing the inverse problem, namely, the estimation of $d_n^{\mathcal{E}}$ from the available information (i.e., $D^{i,\mathcal{E}}$ and K^i). From Eqs. (10) and (16), we obtain

$$\sum_{n\geqslant 0} \phi_{i,n}^{-1}\big(d_n^{\mathcal{E}}\big) \geqslant D^{i,\mathcal{E}} \geqslant \sum_{n\geqslant 0} \Phi_{i,n}^{-1}\big(d_n^{\mathcal{E}}\big); \quad (17)$$

see Fig. 4 for an example.

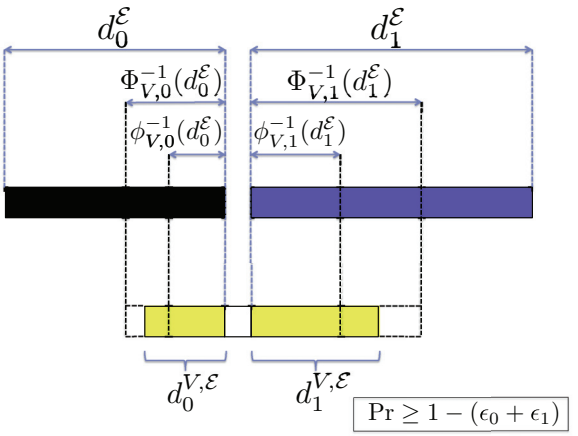

FIG. 4. (Color online) Upper and lower bounds on $d_0^{V,\mathcal{E}} + d_1^{V,\mathcal{E}} \approx d^{V,\mathcal{E}}$. The yellow (bottom) blocks represent the number of pulses from source V with $n = 0$ and $n = 1$. The black (upper left) and blue (upper right) blocks represent the total number of pulses with $n = 0$ and $n = 1$, respectively. The confidence level for this case is not smaller than $1 - (\epsilon_0 + \epsilon_1)$.

Next, our estimation procedure executes a program to obtain $d_0^{\mathcal{E}*}$ and $d_1^{\mathcal{E}*}$, the corresponding smallest values of $d_0^{\mathcal{E}}$ and $d_1^{\mathcal{E}}$, subject to the constraints given by Eqs. (17). From the union bound, the confidence level in such values is $1 - \bar{\epsilon}_{\mathrm{DSP}}$, with

$$\bar{\epsilon}_{\mathrm{DSP}} \leqslant 3 \sum_{n \geqslant 0} \epsilon_n, \tag{18}$$

when three sources are used. Since $f_0^{\mathcal{E}*}$ and $f_1^{\mathcal{E}*}$ are sampled according to a binomial distribution when given $F^{\mathcal{E}}$ (i.e., the preparation and detection basis are random), we obtain

$$f_n^{\mathcal{E}*} = F^{\mathcal{E}} \frac{d_n^{\mathcal{E}*}}{D^{\mathcal{E}}} - c(\bar{\delta}_{\mathrm{DSP}}) \sqrt{F^{\mathcal{E}} \frac{d_n^{\mathcal{E}*}}{D^{\mathcal{E}}} \left(1 - \frac{d_n^{\mathcal{E}*}}{D^{\mathcal{E}}}\right)}, \tag{19}$$

where the constant $c(\bar{\delta}_{\mathrm{DSP}}) \geqslant 0$ can be obtained using Eq. (B3). The overall confidence level for the key rate s is $1 - \epsilon_{\mathrm{DSP}}$, where the security parameter satisfies

$$\epsilon_{\mathrm{DSP}} \leqslant \bar{\epsilon}_{\mathrm{DSP}} + \bar{\delta}_{\mathrm{DSP}}. \tag{20}$$

In the next section we obtain the confidence intervals and levels specifically for our method.

V. CONFIDENCE INTERVALS FOR THE ESTIMATION PROCEDURE

Our method takes ϵ_{DSP} as input and outputs $f_0^{\mathcal{E}*}$ and $f_1^{\mathcal{E}*}$. To satisfy Eq. (20), we can set

$$c(\bar{\delta}_{\mathrm{DSP}}) = 2\sqrt{|\log(\epsilon_{\mathrm{DSP}}/2)|} \tag{21}$$

and

$$\epsilon_n = (\epsilon_{\mathrm{DSP}}/12)(1/2)^n \tag{22}$$

[see Eqs. (B3) and (C1)]. Next, we will find $d_0^{\mathcal{E}*}$ and $d_1^{\mathcal{E}*}$ as required by Eq. (19).

If ϕ depends on $d_n^{i,\mathcal{E}}$ only, the probability that $d_n^{\mathcal{E}}$ is smaller than ϕ is

$$\sum_{d_n^{\mathcal{E}}=0}^{K} \Pr\left(d_n^{\mathcal{E}}\right) \sum_{d_n^{\mathcal{E}} \geqslant d_n^{i,\mathcal{E}} > u_n^i} \Pr\left(d_n^{i,\mathcal{E}} | d_n^{\mathcal{E}}\right) = \frac{\epsilon_n}{2}, \tag{23}$$

with

$$u_n^i = \phi_{i,n}^{-1}\left(d_n^{\mathcal{E}}\right).$$

When given $d_n^{\mathcal{E}}$, the random variable $d_n^{i,\mathcal{E}}$ is sampled according to Eq. (8). From Chernoff's bound (Appendix B)

$$\epsilon_n \leqslant 2 \max_{0 \leqslant d_n^{\mathcal{E}} \leqslant K} \exp\left\{-\frac{\left(u_n^i - q_n^i d_n^{\mathcal{E}}\right)^2}{4q_n^i\left(1-q_n^i\right)d_n^{\mathcal{E}}}\right\}, \tag{24}$$

and we choose the lower bound so that

$$\phi_{i,n}\left(d_n^{i,\mathcal{E}}\right) = \frac{d_n^{i,\mathcal{E}}}{q_n^i} - c_n \frac{1-q_n^i}{2q_n^i}\left[\sqrt{c_n^2 + \frac{4d_n^{i,\mathcal{E}}}{\left(1-q_n^i\right)^2}} - c_n\right], \tag{25}$$

with $c_n \geqslant 0$. The error probability satisfies

$$\epsilon_n \leqslant 2\exp\left\{-c_n^2/4\right\}; \tag{26}$$

see Appendix C. A similar analysis gives the upper bound

$$\Phi_{i,n}\left(d_n^{i,\mathcal{E}}\right) = \frac{d_n^{i,\mathcal{E}}}{q_n^i} + c_n \frac{1-q_n^i}{2q_n^i}\left[\sqrt{c_n^2 + \frac{4d_n^{i,\mathcal{E}}}{\left(1-q_n^i\right)^2}} + c_n\right], \tag{27}$$

with the same confidence level. Then, to satisfy Eq. (22), it suffices to set

$$c_n^2(\epsilon_{\mathrm{DSP}}) = 4|\log(\epsilon_{\mathrm{DSP}}/24) + n\log(1/2)|.$$

To complete the estimation procedure, we invert Eqs. (25) and (27) and obtain

$$\begin{aligned} \sum_{n \geqslant 0} q_n^i d_n^{\mathcal{E}} + c_n(\epsilon_{\mathrm{DSP}})\sqrt{q_n^i\left(1-q_n^i\right)d_n^{\mathcal{E}}} &\geqslant D^{i,\mathcal{E}}, \\ D^{i,\mathcal{E}} \geqslant \sum_{n \geqslant 0} q_n^i d_n^{\mathcal{E}} - c_n(\epsilon_{\mathrm{DSP}})\sqrt{q_n^i\left(1-q_n^i\right)d_n^{\mathcal{E}}}&. \end{aligned} \tag{28}$$

We can then execute a program that finds the minimum values of $d_0^{\mathcal{E}}$ and $d_1^{\mathcal{E}}$ subject to Eqs. (28). For instance, a quadratic program can be used to search $\sqrt{q_n^i d_n^{\mathcal{E}*}}$. Such minimum values will be used in Eqs. (19) and (13) to obtain the key rate.

A technical remark is in order. When $n \to \infty$, $q_n^i(1 - q_n^i) \to 0$ exponentially fast in n. Then, the contribution of large-n terms in Eqs. (28) is negligible. We can set a suitable cutoff $n_{\max} \geqslant n$ in the number of photons per pulse in our analysis, to avoid unnecessary computational overheads in finding $d_0^{\mathcal{E}*}$ and $d_1^{\mathcal{E}*}$, and with an insignificant impact in the estimated values.

VI. CONCLUSIONS

We analyzed general photon-number-splitting attacks and pointed out that previous security analyses on decoy-state protocols for QKD made a strong assumption on the attack. We

provided an estimation procedure that sets a lower bound on the size of the secure, distributed key, with the corresponding confidence levels. Our procedure requires executing a program to find the minimum values of the number of detections due to empty and single-photon pulses, subject to constraints that are determined by the results of the protocol and by the desired security parameter. This results in rigorous security guarantees even if Eve correlates her attack according to the number of photons in the pulse.

We emphasize that our estimation procedure is not unique: Any time that a confidence interval can be set as a function of publicly available information for general attacks, then an estimation procedure is possible. In addition, our choice of confidence intervals and ϵ_n is not essential and could be further optimized to improve the size of the secure key.

ACKNOWLEDGMENTS

We thank Jane Nordholt, Kevin McCabe, Raymond Newell, Charles Peterson, and Stephanie Wehner for discussions. We thank the Laboratory Directed Research and Development (LDRD) Program at Los Alamos National Laboratory for support.

APPENDIX A: PROPERTIES OF Z_i^E

We let $X \in \{0,1,\ldots,K\}$ be a random variable and $f(X)$ the probability distribution. The random variable $Y \in \{0,1,\ldots,K\}$ has the conditional distribution $g(Y|X)$. The probability of Y is $h(Y) = \sum_{X=0}^{K} g(Y|X)f(X)$. Then, it is easy to show that

$$\sigma_Y^2 = E\left[\sigma_{Y|X}^2\right] + \sigma_{E[Y|X]}^2, \tag{A1}$$

where

$$\sigma_Y^2 = \sum_{Y=0}^{K} h(Y)Y^2 - \left(\sum_{Y=0}^{K} h(Y)Y\right)^2$$

is the variance of Y. Also,

$$E[Y|X] = \sum_{Y=0}^{K} g(Y|X)Y$$

is the expected value of Y when given X,

$$\sigma_{E[Y|X]}^2 = \sum_{X=0}^{K} f(X)E[Y|X] - \left(\sum_{X=0}^{K} f(X)E[Y|X]\right)^2$$

is the variance of $E[Y|X]$,

$$\sigma_{Y|X}^2 = \sum_{Y=0}^{K} g(Y|X)Y^2 - \left(\sum_{Y=0}^{K} g(Y|X)Y\right)^2$$

is the variance of Y when given X, and

$$E\left[\sigma_{Y|X}^2\right] = \sum_{X=0}^{K} f(X)\sigma_{Y|X}^2$$

is the expected value of such a variance.

In the attack discussed in Sec. III, K is fixed and the distribution of k_n satisfies

$$E[k_n] = p_n K, \quad \sigma_{k_n}^2 = p_n(1-p_n)K.$$

Next, $d_n^{\mathcal{E}}$ is chosen such that, when given k_n,

$$E\left[d_n^{\mathcal{E}}\middle|k_n\right] = y_n k_n, \quad \sigma_{d_n^{\mathcal{E}}|k_n}^2 = \tau^2 y_n(1-y_n)k_n.$$

It follows that

$$E\left[\sigma_{d_n^{\mathcal{E}}|k_n}^2\right] = \tau^2 y_n(1-y_n)p_n K,$$
$$\sigma_{E\left[d_n^{\mathcal{E}}|k_n\right]}^2 = (y_n)^2\sigma_{k_n}^2 = (y_n)^2 p_n(1-p_n)K.$$

Then,

$$\sigma_{d_n^{\mathcal{E}}}^2 = \tau^2 y_n(1-y_n)p_n K + (y_n)^2 p_n(1-p_n)K.$$

When given $d_n^{\mathcal{E}}$, the distribution for $d_n^{i,\mathcal{E}}$ satisfies

$$E\left[d_n^{i,\mathcal{E}}\middle|d_n^{\mathcal{E}}\right] = q_n^i d_n^{\mathcal{E}}, \quad \sigma_{d_n^{i,\mathcal{E}}|d_n^{\mathcal{E}}}^2 = q_n^i\left(1-q_n^i\right)d_n^{\mathcal{E}}.$$

Therefore,

$$\sigma_{E\left[d_n^{i,\mathcal{E}}\middle|d_n^{\mathcal{E}}\right]}^2 = \left(q_n^i\right)^2\sigma_{d_n^{\mathcal{E}}}^2, \quad E\left[\sigma_{d_n^{i,\mathcal{E}}|d_n^{\mathcal{E}}}^2\right] = q_n^i\left(1-q_n^i\right)y_n p_n K.$$

Also,

$$\begin{aligned}\sigma_{d_n^{i,\mathcal{E}}}^2 &= \left(q_n^i\right)^2\sigma_{d_n^{\mathcal{E}}}^2 + q_n^i\left(1-q_n^i\right)y_n p_n K \\ &= \left(q_n^i\right)^2[\tau^2 y_n(1-y_n)p_n K + (y_n)^2 p_n(1-p_n)K] \\ &\quad + q_n^i\left(1-q_n^i\right)y_n p_n K \\ &= \left[(\tau^2-1)q_n^i(1-y_n) + \left(1-q_n^i y_n p_n\right)\right]q_n^i y_n p_n K.\end{aligned}$$

The first term on the right-hand side vanishes when $\tau = 1$. The second term is

$$\left(1-q_n^i y_n p_n\right)q_n^i y_n p_n K = \left(1-q^i y_n p_n^{\mu^i}\right)q^i y_n p_n^{\mu^i} K,$$

so that

$$\sum_{n\geqslant 0}\sigma_{d_n^{i,\mathcal{E}}}^2 = q^i Y(\mu^i)K - (q^i)^2 K\sum_{n\geqslant 0}\left[y_n p_n^{\mu^i}\right]^2,$$

for $\tau = 1$. Moreover, since $\sum_{n\geqslant 0}[y_n p_n^{\mu^i}]^2 \ll Y(\mu^i) \ll 1$, then

$$\sum_{n\geqslant 0}\sigma_{d_n^{i,\mathcal{E}}}^2 \approx q^i Y(\mu^i)[1-Y(\mu^i)]K,$$

which shows that the case discussed in Sec. II, i.e., the i.i.d. assumption, corresponds to choosing $\tau = 1$ in this case.

APPENDIX B: CHERNOFF BOUND

Chernoff's bound [19] sets a bound on the probabilities of "rare" events as a function of the standard deviation of the corresponding distribution. More precisely, we let $X_1, X_2, \ldots, X_n$ be a set of i.i.d. random variables that satisfy $|X_j| \leqslant 1$ and define $X = \sum_j X_j$. A general version of Chernoff's bound implies

$$\Pr[X > E[X] + c\sigma] \leqslant \exp\{-c^2/4\}, \tag{B1}$$

where $\sigma = n^{1/2}\sqrt{E[(X_j)^2] - (E[X_j])^2}$ is the standard deviation. For the special case of the binomial distribution where

$X_j = 1$ with probability a and $X_j = 0$ otherwise,

$$\sigma = \sqrt{na(1-a)}, \quad E[X] = na,$$

and

$$\Pr[X > k] = I_a(k, n-k+1) \leqslant \exp\{-(k-na)^2/[4a(1-a)n]\}. \quad \text{(B2)}$$

Here, $I_a(k,\, n-k+1)$ is the so-called regularized incomplete beta function. To satisfy $\Pr[X > k] \leqslant \epsilon$, it suffices to choose c such that

$$|c| = 2\sqrt{|\log \epsilon|}. \quad \text{(B3)}$$

APPENDIX C: CALCULATIONS OF ERRORS

If $\epsilon_n \leqslant (\epsilon/12)(1/2)^n$, then

$$\bar{\epsilon} = \sum_i \sum_n \epsilon_n \leqslant 3(\epsilon/12)\cdot 2 = \epsilon/2, \quad \text{(C1)}$$

where we considered that three sources i are involved in the DSP.

Chernoff's bound for the binomial distribution [Eq. (B2)] implies that

$$\epsilon_n \leqslant 2\exp\left\{-\frac{\left(u_n^i - q_n^i d_n^{\mathcal{E}}\right)^2}{4q_n^i\left(1-q_n^i\right)d_n^{\mathcal{E}}}\right\}.$$

If we set

$$u_n^i = \phi_{i,n}^{-1}\left(d_n^{\mathcal{E}}\right) = q_n^i d_n^{\mathcal{E}} + c_n\sqrt{q_n^i\left(1-q_n^i\right)d_n^{\mathcal{E}}}, \quad \text{(C2)}$$

then

$$\epsilon_n \leqslant 2\exp\left\{-c_n^2/4\right\},$$

as in Eq. (26). Replacing u_n^i by $d_n^{i,\mathcal{E}}$ and $d_n^{\mathcal{E}}$ by $\phi_{i,n}(d_n^{i,\mathcal{E}})$ in Eq. (C2), and solving the resulting quadratic equation, we obtain

$$\sqrt{\phi_{i,n}\left(d_n^{i,\mathcal{E}}\right)} = \left[-c_n\sqrt{q_n^i\left(1-q_n^i\right)} + \sqrt{c_n^2 q_n^i\left(1-q_n^i\right) + 4q_n^i d_n^{i,\mathcal{E}}}\right]/\left(2q_n^i\right).$$

That is,

$$\phi_{i,n}\left(d_n^{i,\mathcal{E}}\right) = \frac{d_n^{i,\mathcal{E}}}{q_n^i} + \frac{c_n^2\left(1-q_n^i\right)}{2q_n^i} - \frac{c_n\sqrt{c_n^2\left(1-q_n^i\right)\left[\left(1-q_n^i\right) + 4d_n^{i,\mathcal{E}}\right]}}{2q_n^i},$$

which yields Eq. (25). Changing $c_n \to -c_n$ provides the upper bound without changing ϵ_n, i.e., the confidence level.

[1] S. Wiesner, Sigact News, **15**, 78 (1983).

[2] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India, December 10–12, 1984* (IEEE, Piscataway, NJ, 1984), pp. 175–179.

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

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

[5] D. Mayers, in *Advances in Cryptology—CRYPTO '96*, Lecture Notes in Computer Science (Springer, New York, 1996), Vol. 1109, p. 343.

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

[7] R. Renner, Ph.D. thesis, Swiss Federal Institute of Technology Zurich, 2005.

[8] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1981).

[9] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, J. Cryptology **5**, 3 (1992).

[10] T. Schmitt-Manderbach *et al.*, Phys. Rev. Lett. **98**, 010504 (2007).

[11] D. Rosenberg *et al.*, New J. Phys. **11**, 045009 (2009).

[12] N. Lütkenhaus, Phys. Rev. A **61**, 052304 (2000).

[13] N. Lütkenhaus and M. Jahma, New J. Phys. **4**, 1 (2002).

[14] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, Quantum Inf. Comput. **4**, 325 (2004).

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

[16] H.-K. Lo, X. Ma, and K. Chen, Phys. Rev. Lett. **94**, 230504 (2005).

[17] P. Rice and J. Harrington, arXiv:0901.0013.

[18] Our definition of $D^{i,\mathcal{E}}$ does not imply that the same n-photon pulse created by Alice is detected by Bob, as Eve may be replacing each pulse at her will. $D^{i,\mathcal{E}}$ only gives information about detections corresponding to those clock cycles in which Alice prepared the pulse with source i. Security will follow from the analysis of the bit error rates that will subtract those detections corrupted by Eve as implied by Eq. (13).

[19] W. Hoeffding, J. Am.. Stat. Assoc. **58**, 13 (1963); H. Chernoff, Ann. Probab. **9**, 533 (1981).

[20] We remark that Eq. (5) does not properly regard the problem of inferring a distribution for $Y^{\mathcal{E}}(\mu^i)$ from the known $\frac{D^{i,\mathcal{E}}}{K^i}$, a problem that would require knowledge of the prior distribution of $Y^{\mathcal{E}}(\mu^i)$.

[21] The exact values of $Y(\mu^i)$ are irrelevant as Alice and Bob cannot have exact estimates of the channel parameters such as η, in general. If Eve changes these values slightly, such changes will not be noticed.