


## Using Cascade in quantum key distribution

Devashish Tupkary<sup>\*</sup> and Norbert Lütkenhaus<sup>†</sup>

*Institute for Quantum Computing, and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

 (Received 28 July 2023; revised 6 October 2023; accepted 11 December 2023; published 22 December 2023)

We point out a critical flaw in the analysis of quantum key distribution protocols that employ the two-way error-correction protocol Cascade. Specifically, this flaw stems from an incomplete consideration of all two-way communication that occurs during the Cascade protocol. We present a straightforward and elegant alternative approach that addresses this flaw and produces valid key rates. We exemplify our new approach by comparing its key rates with those generated using older, incorrect approaches, for qubit BB84 and decoy-state BB84 protocols. We show that, in many practically relevant situations, our rectified approach produces the same key rate as older, incorrect approaches. However, in other scenarios, our approach produces valid key rates that are lower, highlighting the importance of properly accounting for all two-way communication during Cascade.

DOI: [10.1103/PhysRevApplied.20.064040](https://doi.org/10.1103/PhysRevApplied.20.064040)

### I. INTRODUCTION

Quantum key distribution (QKD) [1–3] can provide information-theoretic security of secret keys between two communicating parties, Alice and Bob. Since the quantum channel connecting Alice and Bob is not perfect in any practical realization, QKD protocols implement an error-correction step to correct errors in the measurement data collected by Alice and Bob. This involves classical communication between the two parties, and leaks additional information to the eavesdropper Eve, which must be accounted for when calculating the achievable secret key rate. Cascade [4] is one of the most widely used two-way error-correction protocols for QKD. A lot of work has been done optimizing various parameters of the Cascade protocol, such as its block sizes, number of rounds, efficiency, etc. [5–12]. Cascade has also been used in a large number of QKD experiments [13–17].

Our main result is to rectify a flaw in the analysis of QKD protocols using Cascade, which stems from an incomplete consideration of the two-way classical communication during Cascade. We observe that, in past literature, only the communication from Alice to Bob has been accounted for when considering information leakage about the key to Eve. For a rigorous security proof, the

communication from Bob to Alice must also be included when bounding the information leaked to Eve.

We propose a straightforward and elegant alternative approach that produces valid key rates. The main idea is to compute key rates for a protocol that leaks all the communication from Alice to Bob during Cascade, along with all location of errors in Alice and Bob’s raw data to Eve, in the information reconciliation step. We show that this leaks more information to Eve than Cascade, and thus any key rate for such a protocol is a valid key rate for the original protocol that uses Cascade.

We apply our solution to the qubit-based BB84 protocol, and to the polarization-encoded weak coherent pulse (WCP) BB84 with decoy intensities, for a variety of channel models and constraints. We use the numerical framework from Ref. [18] for our calculations. In this work, we restrict our attention to the asymptotic regime for simplicity, where one can assume an independent and identically distributed collective attack without loss of generality [19, 20]. However, our solution can be directly adapted to the analysis of finite-size protocols. This is because many such analyses ultimately involve the optimization of the same objective function [21–24] (with different constraints), and our approach modifies only the objective function.

This paper is organized as follows. In Sec. II we explain the steps in a generic QKD protocol and explain the Cascade protocol briefly. In Sec. III we explain the problem with past analysis of QKD protocols using Cascade, and present our arguments for correcting it. We also review the numerical framework that we used to compute key rates in this work. In Secs. IV and V we apply our solution to the BB84 protocol implemented using qubits, and using WCP states with decoy intensities, respectively. In Sec. VI we present concluding remarks.

<sup>\*</sup>tupkary.devashish@gmail.com

<sup>†</sup>nlutkenhaus.office@uwaterloo.ca

*Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.*

## II. BACKGROUND

### A. Protocol description

In this subsection, we give a description of the asymptotic formulation of a typical QKD protocol that can use Cascade in the information reconciliation step.

(1) *Quantum phase.* In an entanglement-based protocol, Alice and Bob receive states from a source and perform measurements on them. In a prepare-and-measure protocol, Alice prepares and sends signals to Bob, who measures them. The security analysis of a prepare-and-measure scheme can be reduced to that of an entanglement-based scheme using the source-replacement scheme [25].

(2) *Acceptance test (parameter estimation).* Alice and Bob announce the measurements obtained, and signals sent, for a small fraction of signals. They then perform a test to decide whether to abort or continue the protocol. This step is modeled as Alice and Bob performing some measurements given by positive operator-valued measures (POVMs)  $\{\Gamma_k\}$ , obtaining expectation values  $\{\gamma_k\}$ . The POVMs and expectation values depend on whether the protocol implements “fine-graining” or “coarse-graining” during the acceptance test [26], and the exact nature of the coarse-graining.

(3) *Classical processing.* For the remaining signals, Alice and Bob perform some blockwise processing of data. This involves operations such as public announcements and sifting to remove unwanted signals. Alice then implements a key map that maps her local data and the information exchanged in the blockwise processing to her raw key.

(4) *Error correction and verification.* Alice and Bob implement error correction by exchanging classical information. Cascade can be used in this step. Alice and Bob then compare a randomly chosen hash of their raw keys for error verification, and abort the protocol if the hashes do not match.

(5) *Privacy amplification.* Alice and Bob choose a common two-universal hash function and apply it to their raw keys to generate their final secret key.

If  $\tilde{A}$  and  $\tilde{B}$  denote announcements made by Alice and Bob in the blockwise processing step,  $Z$  denotes the result of the key map implemented by Alice, and  $E$  is Eve’s quantum system, then the key rate is given by [19,27]

$$R = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} S(Z | E\tilde{A}\tilde{B}) - p_{\text{pass}}\delta_{\text{leak}}. \quad (1)$$

Here the minimization is over all states  $\rho$  belonging to  $\mathcal{S}(\tilde{\gamma}) = \{\rho \in H_+ \mid \text{Tr}(\Gamma_k\rho) = \gamma_k\}$  where  $H_+$  denotes positive semidefinite operators,  $p_{\text{pass}}$  denotes the probability of the signal to pass sifting, and  $\delta_{\text{leak}}$  is the number of bits used during error correction per bit of raw key.

### B. Cascade

In this subsection we briefly describe the error-correction protocol Cascade [4]. Cascade is a simple and efficient error-correction protocol, and its principal limitation is the requirement for highly interactive communications, as compared to approaches such as low-density parity-check (LDPC) codes (which suffer from a high computational cost in iterative decoding) [28]. To understand Cascade, we first look at a subprotocol called BINARY, which corrects a single error in bit strings that contain an odd number of errors.

(a) *BINARY.* If bit strings  $X$  and  $Y$  have odd numbers of errors, then Alice divides her string into halves and sends the parity of the first half to Bob. Bob divides his string the same way, and announces whether the parity of the first half is wrong, or the parity of the second half is wrong. Alice and Bob repeat the operation on the half whose parity was wrong.

(b) The process terminates when Alice reveals the single bit that contains an error, and Bob corrects that error.

(c) This process involves sending  $\approx \log(k)$  bits from Alice to Bob, and  $\approx \log(k)$  bits from Bob to Alice, where  $k$  is the length of the strings  $X$  and  $Y$ . The process corrects one error.

**Cascade.** The Cascade protocol consists of several passes and proceeds as follows.

(1) Alice and Bob divide their bit strings  $X_1 \cdots X_N$  and  $Y_1 \cdots Y_N$ , where  $N$  is the total number of sifted bits, into blocks of size  $k_1$ . In pass 1, Alice and Bob reveal the parity of each block to determine the blocks with an odd number of errors. For each block with an odd number of errors, Alice and Bob run BINARY to correct one error. At the end of pass 1, all blocks have an even number of errors.

(2) In any pass  $i \geq 1$ , Alice and Bob choose a block size  $k_i$  and random function  $f_i : [1 \cdots N] \rightarrow [1 \cdots N/k_i]$ , which assigns each bit to a block in round  $i$ . The bits whose position belongs to  $K_j^i = \{l \mid f_i(l) = j\}$  form the  $j$ th block in the  $i$ th round.

(3) Alice sends the parity of each block  $P_{(A,i,j)} = \bigoplus_{l \in K_j^i} X_l$  to Bob, who computes his parity for the same block and announces it. For each block where  $P_{(A,i,j)} \neq P_{(B,i,j)}$ , Alice and Bob perform the following operations.

(a) Alice and Bob perform BINARY on the block defined by  $K_j^i$  and correct one error, say at position  $l$ . Now, all blocks in previous rounds which contained  $l$  have an odd number of errors. In this way, a single error corrected in each block in later rounds leads to the identification of several error-containing blocks in earlier rounds. Let the set of such blocks be  $\mathcal{K}$ .

(b) Alice and Bob choose the smallest block from  $\mathcal{K}$  and run BINARY to correct one error. They again compute

the set of blocks containing an odd number of errors  $\mathcal{K}$ . This process is repeated until  $\mathcal{K}$  contains no blocks.

(4) At the end of pass  $i$ , all blocks generated in all rounds contain an even number of errors. Alice and Bob then move to the next pass.

*Remark 1.*—The main ingredient of Cascade that we will use is the fact that, for every parity bit Alice sends to Bob, Bob sends the corresponding parity bit to Alice. There are several variants of the Cascade protocol, which vary in the manner in which blocks are created, block sizes used, and number of passes. Such variations do not change the fact that Bob announces the same set of parities as Alice, and thus our claims will hold for all such variants.

We note that the details of the blocks generated in a given pass have to be communicated between Alice and Bob. However, the blocks are generated randomly and independent of the QKD protocol. Therefore, the act of communicating the details of these blocks does not provide any additional information to Eve about the key [29].

### III. USING CASCADE IN QKD PROTOCOLS

#### A. The problem

In the original proposal for Cascade [4], an analytical upper bound  $\delta_{\text{leak}}^A$  on the number of bits sent from Alice to Bob per bit of raw key is obtained. In an actual experiment, an upper bound  $\delta_{\text{leak}}^A$  can also be chosen empirically by running multiple iterations of Cascade for the expected error rate. For the purposes of this work, it does not matter how  $\delta_{\text{leak}}^A$  is obtained. For convenience, we will denote the upper bound as  $\delta_{\text{leak}}^A = fh(e)$ , where  $e$  is the error rate in the raw key,  $h$  is the binary entropy function, and  $f$  is a number that denotes the efficiency. Typical values of  $f$  for Cascade are between 1 and 1.5, and can be found in Refs. [4,7,8,10].

The original Cascade paper [4] provides only an upper bound on the number of bits sent from Alice to Bob, i.e.,  $\delta_{\text{leak}}^A$ , and defines the “efficiency” of Cascade as the ratio of the actual number of bits per signal sent from Alice to Bob, and  $h(e)$ , where  $e$  is the error rate and  $h$  is the binary entropy function.

Therefore, it has been erroneously assumed that  $\delta_{\text{leak}}^A$  is the true value of  $\delta_{\text{leak}}$  in Eq. (1) when Cascade is used in QKD. It is assumed incorrectly that

$$R_{\text{incorrect}} = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} S(Z | E\tilde{A}\tilde{B}) - p_{\text{pass}}\delta_{\text{leak}}^A \quad (2)$$

is the expression for the key rate. However, *all classical communication* must be assumed to be known to Eve, and the above equation does not account for the communication from Bob to Alice during Cascade. In fact, since Bob’s data are correlated with those of Alice, it is entirely possible for Bob’s communication to leak additional information about Alice’s raw key to Eve.

#### B. A naive approach

One naive approach to fix Eq. (2) is to include  $\delta_{\text{leak}}^B$ , an upper bound on the number of bits leaked during the Bob-to-Alice communication, in  $\delta_{\text{leak}}$  in Eq. (1). Therefore, a naive but correct expression for key rate would be

$$R_{\text{naive}} = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} S(Z | E\tilde{A}\tilde{B}) - p_{\text{pass}}(\delta_{\text{leak}}^A + \delta_{\text{leak}}^B). \quad (3)$$

Here  $\delta_{\text{leak}}^A$  can be replaced with  $fh(e)$ . Cascade requires Bob to send a bit to Alice for every bit Alice sends to Bob. Therefore,  $\delta_{\text{leak}}^B = \delta_{\text{leak}}^A$ , leading to  $(\delta_{\text{leak}}^A + \delta_{\text{leak}}^B) = 2fh(e)$ , which doubles the cost of error correction. Using this value in  $\delta_{\text{leak}}$  in Eq. (3) will yield valid key rates. However, the values obtained will be far worse than those obtained for any one-way error-correction protocol, therefore making Cascade typically unsuitable for QKD.

#### C. Our solution

We show that one can do better than Eq. (3). Recalling Remark 1, we note that the communication from Bob to Alice can be computed from two pieces of information: (1) the communication from Alice to Bob, and (2) the knowledge of the location of errors  $W_i = X_i \oplus Y_i$  for each bit of Alice and Bob’s data. This is because, for any parity bit  $P_{(A,i,j)}$  sent by Alice during Cascade, Bob sends a bit  $P_{(B,i,j)}$  that is the parity of the *same set of bits* of his data. Therefore,  $P_{(A,i,j)} \oplus P_{(B,i,j)} = \sum_{l \in K_j^i} X_l \oplus Y_l = \sum_{l \in K_j^i} W_l$ , where we recall that  $K_j^i$  is the set of positions of the bits that made up the  $j$ th block in the  $i$ th pass of Cascade.

This property implies that a modified protocol that leaks all communication from Alice to Bob, and additionally leaks  $W$ , *can only leak more (or equal) information* to Eve than Cascade. Thus, to lower-bound the key rate for a QKD protocol using Cascade, we can lower-bound the key rate for the QKD protocol that announces  $W$ , and only involves the Alice-to-Bob part of the communication from Cascade. Therefore, we can compute

$$R = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} S(Z | EW\tilde{A}\tilde{B}) - p_{\text{pass}}\delta_{\text{leak}}^A \quad (4)$$

as a valid key rate for any QKD protocol using Cascade.

We note that Eq. (2) will always produce a key rate that is greater than or equal to that produced from Eq. (4), since  $S(Z | E\tilde{A}\tilde{B}) \geq S(Z | EW\tilde{A}\tilde{B})$  from subadditivity. We show that both the equality and inequality can occur, therefore proving that using Eq. (2) can produce key rates that are not justified. In some cases, we obtain equality, which indicates that announcing  $W$  gives Eve no new information. In such cases, although the valid key rate does not change, the argument that properly addresses the communication from Bob to Alice is lacking in the literature and is provided by this work.

Note that a valid key rate for *any* QKD protocol involving Cascade can be obtained by considering the additional announcement of the location of errors along with the Alice-to-Bob communication, as explained above. However, depending on the choice of proof technique, analyzing the protocol with the additional announcement may or may not be straightforward, and should be done carefully. In this work, we demonstrate our solution in the case of proof techniques that analyze single-round entropic quantities.

### D. Computing key rates

We are interested in the difference between the incorrect key rate from Eq. (2) and the key rate from our proposed approach from Eq. (4). Therefore we compute two quantities,  $F = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} S(Z | E\tilde{A}\tilde{B})$  and  $F' = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} S(Z | EW\tilde{A}\tilde{B})$ .

We use the numerical framework from Ref. [18] to perform these optimizations. This framework equivalently describes the steps in the QKD protocol via Kraus operators  $\{K_i\}$ , which represent measurements, announcements, and sifting done by Alice and Bob, and  $\{Z_j\}$ , which implement a pinching channel on the key register. Our solution, which requires the computation of  $F'$  instead of  $F$ , can be easily implemented by a suitable change in the Kraus operators for the optimization problem for  $F$ .

The numerical framework equivalently describes the optimization problem for  $F$  as

$$F = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} f(\rho), \quad (5)$$

where

$$\begin{aligned} f(\rho) &= D(\mathcal{G}(\rho) || \mathcal{Z}(\mathcal{G}(\rho))), \\ \mathcal{G}(\rho) &= \sum_i K_i \rho K_i^\dagger, \\ \mathcal{Z}(\mathcal{G}(\rho)) &= \sum_j Z_j \mathcal{G}(\rho) Z_j^\dagger, \end{aligned} \quad (6)$$

and where  $D(X || Y) = \text{Tr}(X \log(X)) - \text{Tr}(X \log(Y))$  is the quantum relative entropy with  $\log$  as the matrix logarithm.

Let  $\alpha$  and  $\beta$  be announcements (such as basis choice) made by Alice and Bob, and let Alice and Bob's POVMs be given by  $P^A = \{P^A_{(\alpha,x)}\}$  and  $P^B = \{P^B_{(\beta,y)}\}$ , where  $x$  and  $y$  denote bits that represent measurement outcomes. Let  $\mathbf{A}$  be the set of announcements  $(\alpha, \beta)$  that are kept after sifting. Furthermore, let  $r(\alpha, \beta, x)$  be the key map that Alice implements. Then, the Kraus operators for Eq. (2)

are given by

$$\begin{aligned} K_{\alpha,\beta} &= \sum_{x,y} |r(\alpha, \beta, x)\rangle_Z \otimes \sqrt{P^A_{(\alpha,x)} \otimes P^B_{(\beta,y)}} \\ &\otimes |x\rangle_X \otimes |y\rangle_Y \otimes |\alpha, \beta\rangle_{\tilde{A}\tilde{B}}, \end{aligned} \quad (7)$$

and the set of operators generating the  $\mathcal{G}$  map is given by  $\{K_i\} = \{K_{\alpha,\beta} | (\alpha, \beta) \in \mathbf{A}\}$  [18]. The  $\mathcal{Z}$  map is implemented by operators  $\{Z_i\}$  given by  $Z_i = |i\rangle \langle i|_Z \otimes I_{ABXY\tilde{A}\tilde{B}}$ . Notice that the output state  $\mathcal{G}(\rho)$  is classical in  $\alpha, \beta$ , which reflects the fact that the basis choices are announced and known to Eve.

To compute  $F'$ , we must include an additional announcement that announces  $w = x \oplus y$ . This is implemented by

$$\begin{aligned} K'_{\alpha,\beta,w} &= \sum_w \sum_{\substack{x,y \\ x \oplus y = w}} |r(\alpha, \beta, x)\rangle_Z \otimes \sqrt{P^A_{(\alpha,x)} \otimes P^B_{(\beta,y)}} \\ &\otimes |x\rangle_X \otimes |y\rangle_Y \otimes |\alpha, \beta\rangle_{\tilde{A}\tilde{B}} \otimes |w\rangle_W, \end{aligned} \quad (8)$$

where the set of operators generating the new  $\mathcal{G}'$  map can be given by  $\{K'_i\} = \{K'_{\alpha,\beta,w} | (\alpha, \beta) \in \mathbf{A}\}$ . The  $\mathcal{Z}'$  map is implemented by  $\{Z'_i\}$  given by  $Z'_i = |i\rangle \langle i|_Z \otimes I_{ABXY\tilde{A}\tilde{B}W}$ .

In the remainder of this paper, we compute both  $F = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} f(\rho)$  and  $F' = \min_{\rho \in \mathcal{S}(\tilde{\gamma})} f'(\rho)$  for the various implementations of the BB84 protocol. If we find that  $F = F'$ , then this indicates that the previous analysis of Cascade is wrong but gives correct answers. In this case, Eqs. (2) and (4) will give identical key rates. If we find that  $F > F'$ , then this indicates that the previous analysis was wrong and gave incorrect answers. The difference between the key rates obtained from Eqs. (2) and (4) is equal to  $F - F'$ . Note that  $F - F'$  is the difference in the key rates measured in terms of secure key bits per signal sent. The difference in key rates in terms of secure key bits per second is given by  $R_r(F - F')$ , where  $R_r$  is the repetition rate.

Note that the above formulation applies to situations where Alice and Bob generate a bit string from their measurements ( $x$ ,  $y$ , and  $x \oplus y$  are bits). Events such as no-detection either need to be discarded during sifting, or should be mapped to bits. This assumption is necessary to use Cascade, since it is a protocol that corrects errors in *bit strings*. We also remark that, since many finite-size key rate analyses involve the optimization of the same objective function ( $F$ ) over different constraints, our solution can be easily applied to such finite-size analyses as well, by simply changing  $F$  to  $F'$ .

## IV. QUBIT BB84

### A. Protocol description

In this section, we present our results for the standard qubit-based BB84 protocol, where Alice prepares each of

TABLE I. Format of fine-grained statistics obtained for qubit BB84. The rows denote the state sent by Alice, and the columns denote the measurement outcome measured by Bob. The  $H$  and  $+$  correspond to measurement outcome 0, while the  $V$  and  $-$  correspond to measurement outcome 1; and  $\gamma_{x,y}$  denotes the probability of Alice obtaining outcome  $x$  and Bob obtaining outcome  $y$ .

		Bob measures			
		H	V	+	-
Alice sends	H	$\gamma_{HH}$	$\gamma_{HV}$	$\gamma_{H+}$	$\gamma_{H-}$
	V	$\gamma_{VH}$	$\gamma_{VV}$	$\gamma_{V+}$	$\gamma_{V-}$
	+	$\gamma_{+H}$	$\gamma_{+V}$	$\gamma_{++}$	$\gamma_{+-}$
	-	$\gamma_{-H}$	$\gamma_{-V}$	$\gamma_{-+}$	$\gamma_{--}$

the four signal states  $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$  with equal probability, and Bob chooses the  $Z$  or  $X$  basis with equal probability. Alice and Bob then implement an acceptance test on their observed statistics. If the protocol accepts, Alice and Bob announce their basis, and throw away signals where they measured in different bases. Alice maps her measurement outcome to the raw key, and then proceeds to perform error correction (Cascade) and privacy amplification. For the descriptions of the exact Kraus operators of the protocol, we refer the reader to Appendix A 1.

Alice and Bob obtain statistics shown in Table I during the acceptance test. There are a variety of ways they can use these statistics to perform the acceptance test. We use the phrase “fine-grained” constraints to refer to the case where all the entries in Table I are used for the acceptance test, and therefore in the constraints for  $\mathcal{S}(\vec{\gamma})$ . We use “sifted fine-grained” to refer to the case where only the entries marked in red are used. We use “coarse-grained” constraints to refer to the case where only the (unnormalized) quantum bit error rate (QBER) and gain constraints for each basis (given by  $Q_Z = \gamma_{HV} + \gamma_{VH}$  and  $Q_X = \gamma_{+-} + \gamma_{-+}$ , and  $\text{gain}_Z = \gamma_{HH} + \gamma_{HV} + \gamma_{VH} + \gamma_{VV}$  and  $\text{gain}_X = \gamma_{++} + \gamma_{+-} + \gamma_{-+} + \gamma_{--}$ ) are used. The gain sets the constraints on the probability of choosing each basis for measurement, while the QBER in each basis sets the constraints on the observed error rate. We note that this is a departure from the nomenclature of Ref. [26], where the “coarse-grained” case refers to the “sifted fine-grained” case as defined above. Additionally, we use the constraints from source replacement that characterize Alice’s system for prepare-and-measure protocols.

We consider a channel with misalignment and depolarization to compute statistics in Table I, which is described in Appendix A 1. We also consider a channel model that includes a “replacement” channel  $\Phi_{\text{replace}}$ , which replaces the state leaving Alice’s lab with the fixed signal state  $|0\rangle$  ( $|H\rangle$ ) with probability  $\lambda = 0.2$ . The output of the replacement channel is then sent through a channel with misalignment and depolarization. The replacement channel is

interesting because it breaks symmetries in the observed statistics. If the replacement channel is also included in the channel model, then the new statistics can be obtained by replacing each row  $\vec{\gamma}_i$  of Table I by  $(1 - \lambda)\vec{\gamma}_i + \lambda\vec{\gamma}_H$  (since Alice sends each state with equal probability).

## B. Reduction to Bell-diagonal states

In certain cases, the optimization of  $f(\rho)$  over all states  $\rho$  in  $\mathcal{S}(\vec{\gamma})$  can be reduced to that over all Bell-diagonal states in  $\mathcal{S}(\vec{\gamma})$ , denoted by  $\mathcal{S}_{\text{Bell}}(\vec{\gamma})$ . That is, it can be shown that

$$\min_{\rho \in \mathcal{S}_{\text{Bell}}(\vec{\gamma})} f(\rho) = \min_{\rho \in \mathcal{S}(\vec{\gamma})} f(\rho). \quad (9)$$

For Bell-diagonal states shared between Alice and Bob, Eve’s state  $E$  is always block-diagonal in the parities  $\mathcal{W}$  (see Appendix B), and therefore the additional announcement  $\mathcal{W}$  gives Eve no new information. In such cases,  $F = F'$ .

There are several such arguments in the literature, which are identical at their core, but differ only in details of the protocols (such as type of constraints, number of bases used for key generation, and type of classical processing). Ref. [30] proves Eq. (9) for the case where  $f(\rho)$  is the total key rate including the information leakage term, and the states are constrained in  $\mathcal{S}$  only by the average QBER over all bases. The analysis is done for  $d$ -dimensional systems in general. Ref. [31] proves Eq. (9) for BB84 and six-state protocols, where  $f(\rho)$  is the uncertainty of Eve about the raw key (with a modified classical processing), and the states are constrained in  $\mathcal{S}$  by each individual QBER, but only the  $Z$  basis is used for key generation. Ref. [32] generalizes this to a wider variety of classical processing in  $f(\rho)$ , while still constraining  $\mathcal{S}$  with separate QBERs, but using only the  $Z$  basis for key generation. In this work, we will attempt to present a coherent unified picture of all such arguments for the convenience of the reader. We also point out how symmetry in *observed values* can help in proving the reduction to Bell-diagonal states.

We start by defining the “twirling map” [33] as

$$\mathcal{T}(\rho) = \frac{1}{4} \sum_{i=1}^4 \rho_i = \frac{1}{4} \sum_{i=1}^4 (\sigma_i \otimes \sigma_i) \rho (\sigma_i \otimes \sigma_i)^\dagger, \quad (10)$$

where  $\sigma_i$  for  $i \in \{0, 1, 2, 3\}$  denotes the identity and the Pauli  $X$ ,  $Y$ , and  $Z$  operators. The map  $\mathcal{T}(\rho)$  is often referred to as the “twirled” state, and can be shown to be always Bell-diagonal [30,33]. The proof of Eq. (9) now proceeds in two steps.

*Step 1:* It is shown that

$$f(\mathcal{T}(\rho)) \leq f(\rho) \quad \forall \rho. \quad (11)$$

*Step 2:* It is shown that

$$\rho \in \mathcal{S} \implies \mathcal{T}(\rho) \in \mathcal{S}_{\text{Bell}}. \quad (12)$$

The proof of Eq. (9) then proceeds as follows. Clearly  $\min_{\rho \in \mathcal{S}_{\text{Bell}}(\vec{\gamma})} f(\rho) \geq \min_{\rho \in \mathcal{S}(\vec{\gamma})} f(\rho)$ , since  $\mathcal{S}_{\text{Bell}}(\vec{\gamma}) \subseteq \mathcal{S}(\vec{\gamma})$ . To show the other direction of the inequality, let  $\rho^*$  be the state that achieves the minimization on the right-hand side of Eq. (9). Then, from Eq. (11), we obtain  $f(\mathcal{T}(\rho^*)) \leq f(\rho^*)$ . From Eq. (12), we know that  $\mathcal{T}(\rho^*) \in \mathcal{S}_{\text{Bell}}(\vec{\gamma})$ . Therefore,  $\min_{\rho \in \mathcal{S}_{\text{Bell}}(\vec{\gamma})} f(\rho) \leq f(\mathcal{T}(\rho^*)) \leq \min_{\rho \in \mathcal{S}(\vec{\gamma})} f(\rho)$ .

Thus, to obtain Eq. (9) for a protocol of interest, one must show the validity of Eqs. (11) and (12). We prove that Eq. (11) holds for qubit protocols where key generation is done in all the  $X$ ,  $Z$  (and, if applicable,  $Y$ ) bases in Appendix C. Thus, to reduce the optimization to Bell-diagonal states and obtain  $F = F'$ , we now only need to check the validity of Eq. (12). This has to be considered separately for every choice of constraints, and observed values, and is done in the next section.

### C. Results

We numerically check the difference between  $F$  and  $F'$ . The results are summarized in Table II. Since the numerical method is capable of producing both an upper bound and a lower bound for  $F$  and  $F'$ , it is straightforward to determine when  $F > F'$ . We claim  $F = F'$  when the bounds for  $F$  and  $F'$  overlap. In some cases,  $F = F'$  can be analytically argued, by proving the validity of Eq. (12) (see Sec. IV B), as we do below.

To check the validity of Eq. (12), we need to look at the constraints that define  $\mathcal{S}(\vec{\gamma})$ . There are two types of constraints. First, we have the constraint obtained from the source-replacement scheme, which is of the form  $\text{Tr}_B(\rho) = \sigma_A$  and represents the fact that, in prepare-and-measure protocols, Alice's state is known and never leaves her lab. For the qubit BB84 protocol, one can take  $\sigma_A = I_A/2$  (see Appendix A 1). Since  $\text{Tr}_B(\rho_{AB}) = I_A/2$  is always true for a Bell-diagonal state, this constraint is always satisfied by  $\mathcal{T}(\rho)$ .

The remaining constraints are obtained from the acceptance test, and are of the form  $\text{Tr}(\Gamma_k \rho) = \gamma_k$ . Thus, to check the validity of Eq. (12), we check whether  $\text{Tr}(\Gamma_k \mathcal{T}(\rho)) = \text{Tr}(\mathcal{T}^\dagger(\Gamma_k) \rho) = \gamma_k$  for all  $\rho$ .

(a) *Coarse-grained statistics.* In this case,  $\Gamma_k$  is the POVM element corresponding to QBER and gain in each basis. It is therefore easy to check with a simple calculation that  $\mathcal{T}^\dagger(\Gamma_k) = \Gamma_k$ , which implies  $\text{Tr}(\Gamma_k \mathcal{T}(\rho)) = \text{Tr}(\Gamma_k \rho)$ . Thus, for coarse-grained constraints, the state shared between Alice and Bob can be assumed to be Bell-diagonal, and  $F = F'$ . This explains the coarse-grained row in Table II.

(b) *Sifted fine-grained statistics.* Let us turn to the case of ‘‘sifted fine-grained’’ constraints. For the  $Z$  basis, let the POVMs that make up the constraints be given by  $\Gamma_{HH}$ ,  $\Gamma_{HV}$ ,  $\Gamma_{VH}$ , and  $\Gamma_{VV}$  (with similar expressions for the  $X$  basis). In general, these POVMs are not invariant under  $\mathcal{T}^\dagger$ , and thus  $\text{Tr}(\Gamma_k \mathcal{T}(\rho)) \neq \text{Tr}(\Gamma_k \rho)$ . However, one can find that

$$\begin{aligned} \mathcal{T}^\dagger(\Gamma_{HH}) &= \mathcal{T}^\dagger(\Gamma_{VV}) = \frac{1}{2}(\Gamma_{HH} + \Gamma_{VV}), \\ \mathcal{T}^\dagger(\Gamma_{HV}) &= \mathcal{T}^\dagger(\Gamma_{VH}) = \frac{1}{2}(\Gamma_{VH} + \Gamma_{HV}), \\ \mathcal{T}^\dagger(\Gamma_{++}) &= \mathcal{T}^\dagger(\Gamma_{--}) = \frac{1}{2}(\Gamma_{++} + \Gamma_{--}), \\ \mathcal{T}^\dagger(\Gamma_{+-}) &= \mathcal{T}^\dagger(\Gamma_{-+}) = \frac{1}{2}(\Gamma_{+-} + \Gamma_{-+}). \end{aligned} \quad (13)$$

Therefore, in this case, one can claim a reduction to Bell-diagonal as long as the statistics obey certain symmetries. That is, if one obtains statistics satisfying  $\gamma_{HH} = \gamma_{VV}$ ,  $\gamma_{HV} = \gamma_{VH}$ ,  $\gamma_{++} = \gamma_{--}$ , and  $\gamma_{+-} = \gamma_{-+}$ , then

$$\begin{aligned} \text{Tr}(\Gamma_k \rho) = \gamma_k &\implies \text{Tr}(\Gamma_k \mathcal{T}(\rho)) = \gamma_k, \\ &\text{even when } \mathcal{T}^\dagger(\Gamma_k) \neq \Gamma_k. \end{aligned} \quad (14)$$

The statistics obey this symmetry when the channel consists of any combination of loss and misalignment, and therefore for these channel models we again obtain  $F = F'$  due to the reduction to Bell-diagonal states. Introducing the additional replacement channel  $\Phi_{\text{replace}}$  destroys this symmetry, and we obtain  $F \neq F'$ . This explains the sifted fine-grained row in Table II.

TABLE II. Relation between  $F$  and  $F'$  for qubit BB84 protocol. Results are based upon the upper and lower bounds to the optimization obtained from the numerical method [18]. All equality cases can be explained by arguing a reduction to Bell-diagonal states.

		Channel			
		Misalignment	Depolarization	Misalignment + depolarization	Misalignment + depolarization + $\Phi_{\text{replace}}$
Constraints	Coarse-grained	=	=	=	=
	Sifted fine-grained	=	=	=	>
	Fine-grained	=	=	>	>

(c) *Fine-grained statistics.* In this case, in addition to Eq. (13), it is possible to show that each POVM in the off-diagonal block of Table I is mapped to the same off-diagonal block by  $\mathcal{T}^\dagger$ . That is,

$$\begin{aligned}\mathcal{T}^\dagger(\Gamma_{H+}) &= \mathcal{T}^\dagger(\Gamma_{H-}) = \mathcal{T}^\dagger(\Gamma_{V+}) = \mathcal{T}^\dagger(\Gamma_{V-}) \\ &= \frac{1}{4}(\Gamma_{H+} + \Gamma_{H-} + \Gamma_{V+} + \Gamma_{V-}), \\ \mathcal{T}^\dagger(\Gamma_{+H}) &= \mathcal{T}^\dagger(\Gamma_{+V}) = \mathcal{T}^\dagger(\Gamma_{-H}) = \mathcal{T}^\dagger(\Gamma_{-V}) \\ &= \frac{1}{4}(\Gamma_{+H} + \Gamma_{+V} + \Gamma_{-H} + \Gamma_{-V}).\end{aligned}\quad (15)$$

That is, if one obtains statistics satisfying  $\gamma_{HH} = \gamma_{VV}$ ,  $\gamma_{HV} = \gamma_{VH}$ ,  $\gamma_{++} = \gamma_{--}$ , and  $\gamma_{+-} = \gamma_{-+}$  along with  $\gamma_{H+} = \gamma_{H-} = \gamma_{V+} = \gamma_{V-}$  and  $\gamma_{+H} = \gamma_{+V} = \gamma_{-H} = \gamma_{-V}$ , then we again can claim that

$$\begin{aligned}\text{Tr}(\Gamma_k \rho) = \gamma_k &\implies \text{Tr}(\Gamma_k \mathcal{T}(\rho)) = \gamma_k, \\ \text{even when } \mathcal{T}^\dagger(\Gamma_k) &\neq \Gamma_k.\end{aligned}\quad (16)$$

This is the case when the channel contains only depolarization.

For only misalignment, it has already been shown that fine-grained constraints allow us to show that Eve factors off and holds a state that is independent of the Alice–Bob state [26]. Since Eve’s quantum system factors off, the  $F = F'$  follows from the fact that  $W$  and  $Z$  are independent random variables for each basis, i.e.,  $S(Z | \tilde{A}\tilde{B}W) = S(Z | \tilde{A}\tilde{B})$ . For the remaining two cases, we find that the optimal values of the two objective functions are unequal, and no reduction to the Bell-diagonal case is possible. This explains the fine-grained row of Table II.

We plot  $F$  and  $F'$  corresponding to the last two columns of Table IV in Figs. 1 and 2. Note that we always find that the fine-grained  $F$  ( $F'$ ) is higher than the sifted fine-grained  $F$  ( $F'$ ), followed by the coarse-grained  $F$  ( $F'$ ), as has already been pointed out in Ref. [26].

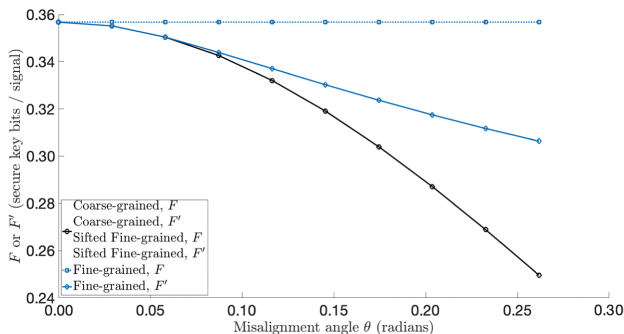


FIG. 1. Plot of  $F$  and  $F'$  for a channel with misalignment and depolarization. We find that  $F = F'$  for coarse-grained and sifted fine-grained constraints, while  $F > F'$  for fine-grained constraints. The plot corresponds to depolarization probability  $q = 0.1$ , and is plotted against the misalignment angle  $\theta$ .

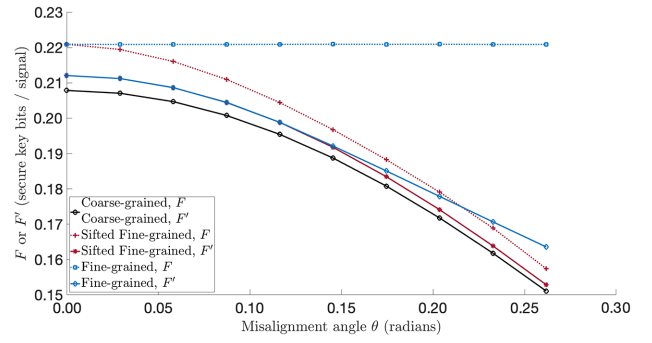


FIG. 2. Plot of  $F$  and  $F'$  for a channel with misalignment, depolarization, and replacement channel. We find that  $F = F'$  for coarse-grained, while  $F > F'$  for sifted fine-grained and fine-grained constraints. The state leaving Alice’s lab is replaced with the signal state corresponding to  $H$ , with probability  $\lambda = 0.2$ . The plot corresponds to depolarization probability  $q = 0.1$ , and is plotted against the misalignment angle  $\theta$ .

## V. WCP DECOY-STATE BB84

In this section, we present results for the WCP decoy-state BB84 protocol [34–40] in the same manner. The key rate calculations, protocol description, channel simulation, and decoy analysis are exactly identical to those from Ref. [26]. Therefore, these aspects will be described only briefly in this work. The only difference lies in the modification of Kraus operators according to Eq. (8), and the inclusion of the replacement channel  $\Phi_{\text{replace}}$  in the channel simulation.

### A. Protocol specification

Alice prepares and sends a phase-randomized weak coherent pulse in the polarization mode  $H, V, A, D$  with equal probability, choosing to use the “signal intensity” with probability close to one, and some “decoy intensities.” Bob implements passive basis choice with equal probability. We use a squashing model on Bob’s side [41] to describe his measurements, and Bob’s squashed POVMs can be found in Appendix A 2. Alice and Bob announce a small fraction of their data, and perform the acceptance test. If the protocol accepts, Alice and Bob announce their basis, and throw away the signals where they measured in different bases, or where Bob got a no-detection event. Alice then maps her measurement outcomes to the raw key, followed by error correction (Cascade) and privacy amplification. For the descriptions of the exact Kraus operators of the protocol, we refer the reader to Appendix A 2.

The fine-grained statistics obtained by Alice and Bob are given in Table III. Again, as in Sec. IV, we use the phrase “fine-grained” constraints to refer to the case where all the entries in Table I are used for the acceptance test, “sifted fine-grained” when only the entries marked in red

TABLE III. Format of fine-grained statistics obtained for decoy-state BB84 protocol. The rows denote the state sent by Alice, and the columns denote the measurement outcome measured by Bob. One such table is obtained for each intensity used by Alice. The  $H$  and  $+$  correspond to measurement outcome 0, while the  $V$  and  $-$  correspond to measurement outcome 1; and  $\gamma_{x,y}^{\mu_i}$  denotes the probability of Alice obtaining outcome  $x$  and Bob obtaining outcome  $y$ , given intensity  $\mu_i$  was used.

		Bob measures				
		H	V	+	-	$\emptyset$
Alice sends	H	$\gamma_{HH}^{\mu_i}$	$\gamma_{HV}^{\mu_i}$	$\gamma_{H+}^{\mu_i}$	$\gamma_{H-}^{\mu_i}$	$\gamma_{H\emptyset}^{\mu_i}$
	V	$\gamma_{VH}^{\mu_i}$	$\gamma_{VV}^{\mu_i}$	$\gamma_{V+}^{\mu_i}$	$\gamma_{V-}^{\mu_i}$	$\gamma_{V\emptyset}^{\mu_i}$
	+	$\gamma_{+H}^{\mu_i}$	$\gamma_{+V}^{\mu_i}$	$\gamma_{++}^{\mu_i}$	$\gamma_{+-}^{\mu_i}$	$\gamma_{+\emptyset}^{\mu_i}$
	-	$\gamma_{-H}^{\mu_i}$	$\gamma_{-V}^{\mu_i}$	$\gamma_{-+}^{\mu_i}$	$\gamma_{--}^{\mu_i}$	$\gamma_{-\emptyset}^{\mu_i}$

are used, and “coarse-grained” constraints when only the (unnormalized) QBER and gain constraints for each basis are used. Additionally, we use the constraints from source replacement that characterize Alice’s system for prepare-and-measure protocols.

The exact manner in which statistics from a channel consisting of misalignment and loss are computed is identical to the procedure described in Appendix C of Ref. [26]. We will not repeat those calculations here. We include an additional “replacement channel”  $\Phi_{\text{replace}}$ , which replaces the state leaving Alice’s lab with the signal state corresponding to  $H$  with probability  $\lambda = 0.2$ . This is interesting since it breaks symmetries in observed statistics. If we include the replacement channel, then each row  $\vec{\gamma}_j^{\mu_i}$  of Table III (computed for loss and misalignment) is replaced by  $(1 - \lambda)\vec{\gamma}_j^{\mu_i} + \lambda\vec{\gamma}_H^{\mu_i}$  (since Alice sends each state with equal probability).

## B. Results

The optimization problem for decoy protocols is solved by obtaining bounds on the single-photon yields as

$$\gamma_{y|x}^{1,L} \leq \gamma_{y|x}^1 \leq \gamma_{y|x}^{1,U}, \quad \forall x, y, \quad (17)$$

where  $\gamma_{y|x}^1$  denotes the probability of Bob obtaining outcome  $y$ , given Alice sent signal  $x$  and 1 photon. One can

then compute lower and upper bounds on  $\gamma_{x,y}^1$ , by using  $\gamma_{x,y}^1 = \Pr(x)\gamma_{y|x}^1$ , where  $\Pr(x)$  denotes the probability of Alice sending signal  $x$ .

The optimization problem [26,35,37,39] is then given by (see Appendix D)

$$F = \min_{\rho \in \mathcal{S}'_1} f(\rho), \quad (18)$$

$$\mathcal{S}'_1 = \{\rho \in H_+ | \gamma_k^{1,L} \leq \text{Tr}(\Gamma_k \rho) \leq \gamma_k^{1,U}, \forall k\},$$

where  $H_+$  denotes positive semidefinite operators,  $\mathcal{S}'_1$  is the set of density operators compatible with observed statistics, and  $k$  depends on the exact nature of coarse-graining.

We numerically compute the difference between  $F = \min_{\rho \in \mathcal{S}'_1} f(\rho)$  and  $F' = \min_{\rho \in \mathcal{S}'_1} f'(\rho)$  for all our channel models, and various types of constraints. The results are summarized in Table IV.

Since after squashing, the single-photon contribution to the objective function involves both Alice and Bob having qubits (or vacuum), our intuition from the qubit BB84 picture can be used to understand the results in Table IV. We believe a more rigorous justification can be made along the same lines as for the qubit case; however, that is not a contribution of this work. For coarse-grained constraints, we expect symmetry arguments to allow us to restrict to Bell-diagonal states, in which case announcing  $W$  provides no new information to Eve. For the “sifted fine-grained” and “fine-grained” cases, we expect symmetry arguments not to work in general, but to allow a restriction to Bell-diagonal states if observations are also symmetric, as seen in Sec. IV. We plot  $F$  and  $F'$  corresponding to the last two columns of Table IV in Figs. 3 and 4.

**Effect of zero-photon contribution.** The above analysis is done for the case where we only keep the single-photon contribution to the key in Eq. (18) (see Eq. (D8)). Let us consider the zero-photon contribution to the key. In this case, note that, since no signal left Alice’s lab, Eve cannot know anything about Alice’s key bit. Therefore, the zero-photon contribution to  $F$  is equal to  $p_{\text{pass}}^{(0)}$ , where  $p_{\text{pass}}^{(0)}$  is the probability of zero-photon event leading to detection and passing sifting.

TABLE IV. Relation between  $F$  and  $F'$  for decoy BB84 protocol. Results are based upon the upper and lower bounds to the optimization obtained from the numerical method [18]. Note that this table is similar to Table II obtained for qubit BB84, suggesting the fact that similar arguments can be made for understanding this table in both cases.

		Channel			
		Loss	Misalignment	Loss + misalignment	Loss + misalignment + $\Phi_{\text{replace}}$
Constraints	Coarse-grained	=	=	=	=
	Sifted fine-grained	=	=	=	>
	Fine-grained	=	=	>	>



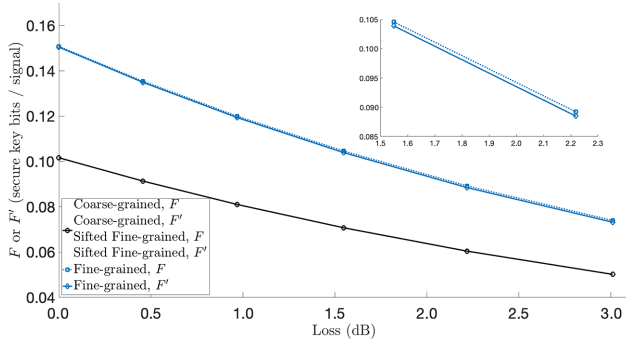


FIG. 3. Plot of  $F$  and  $F'$  for a channel with misalignment and loss. We find that  $F = F'$  for coarse-grained and sifted fine-grained constraints, while  $F > F'$  for fine-grained constraints. The plot corresponds to a misalignment angle  $\theta$  given by  $\sin^2(\theta) = 0.06$ , and three intensities  $\mu_1 = 0.5$ ,  $\mu_2 = 0.1$ , and  $\mu_3 = 0.001$ , with the first intensity used to generate the key.

Moreover, if Alice sends zero photons, the state giving rise to Bob's detection must be assumed to be known to Eve. Therefore, Eve has perfect knowledge of Bob's data. In this case, if  $W$  is announced, Eve has perfect knowledge of Alice's data as well. Therefore, the zero-photon contribution to  $F'$  is *always zero*.

Therefore, in this case  $F > F'$  always, regardless of the type of constraints used.

## VI. CONCLUSION

In this work, we pointed out a critical flaw in the analysis of QKD protocols using Cascade, which stems from an improper consideration of the classical communication during Cascade. This leads to the computation of secret key rates that are not justified. We proposed a simple and

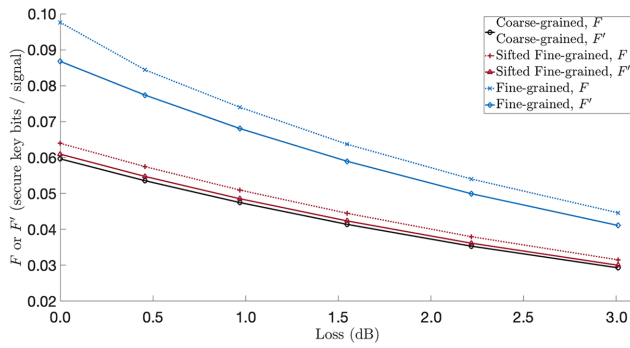


FIG. 4. Plot of  $F$  and  $F'$  for a channel with misalignment, loss, and replacement channel. We find that  $F = F'$  for coarse-grained, while  $F > F'$  for sifted fine-grained and fine-grained constraints. The state leaving Alice's lab is replaced with the signal state corresponding to  $H$  with probability  $\lambda = 0.2$ . The plot corresponds to a misalignment angle  $\theta$  given by  $\sin^2(\theta) = 0.06$ , and three intensities  $\mu_1 = 0.5$ ,  $\mu_2 = 0.1$ , and  $\mu_3 = 0.001$ , with the first intensity used to generate the key.

elegant fix, involving the construction of a convenient virtual protocol that cannot leak less information to Eve than the one using Cascade. Therefore, its key rate can be safely used in any protocol using Cascade. Our approach is easy and straightforward to implement in the numerical framework of Ref. [26]. We applied our solution to various implementations of the BB84 protocol, and compared our results with those of earlier, incorrect approaches. In many cases, we found that the numerical value of the key rate does not change, indicating that the communication from Bob to Alice does not leak additional information to Eve. A number of such cases were shown to arise due to symmetries in the protocol, and in the observed statistics. All code used in this work will be made available soon.

## ACKNOWLEDGMENTS

We would like to thank Lars Kamin and Shlok Nahar for giving useful feedback on early drafts of this work. We would like to thank John Burniston and Scott Johnston for help with debugging code. We would like to thank Louis Salvail for helpful discussions on the Cascade protocol, and for sending us his PhD thesis. This work was funded by the NSERC Discovery Grant, and was conducted at the Institute for Quantum Computing, University of Waterloo, which is funded by the Government of Canada through ISED.

## APPENDIX A: PROTOCOL DESCRIPTIONS

### 1. Qubit BB84

Using the source-replacement scheme [25], the protocol can be equivalently described as Alice creating the Bell state  $|\psi\rangle_{AA'} = |\phi_+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ , and sending  $A'$  to Bob. We model misalignment as a rotation of angle  $\theta$  about the  $Y$  axis on  $A'$ , with

$$U(\theta) = I_A \otimes \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad (\text{A1})$$

$$\mathcal{E}_{\text{misalign}}(\rho) = U(\theta)\rho U(\theta)^\dagger.$$

Depolarization is modeled as a map

$$\mathcal{E}_{\text{depol}}(\rho) = (1 - q)(\rho) + q \text{Tr}_{A'}(\rho) \otimes \frac{I_B}{2}. \quad (\text{A2})$$

The state on which statistics are computed is given by  $\rho_{AB} = \mathcal{E}_{\text{depol}}(\mathcal{E}_{\text{misalign}}(|\phi_+\rangle \langle\phi_+|))$ . The entries in Table I can be computed via  $\gamma_i = \text{Tr}(\Gamma_i \rho_{AB})$ .

Both Alice and Bob perform measurements on qubit systems, and their POVMs are given by  $\{P_{(Z,0)} = p_z |0\rangle \langle 0|, P_{(Z,1)} = p_z |1\rangle \langle 1|, P_{(X,0)} = p_x |+\rangle \langle +|, P_{(X,1)} = p_x |-\rangle \langle -|\}$ , with  $p_z = p_x = 1/2$ . In addition, Alice implements the key map by simply copying the measurement outcome to the key register. From the discussion in

Appendix A of Ref. [42], we can remove certain registers created by the generic form of the Kraus operators in Eq. (7). In particular, we do not need to consider the registers that store Alice and Bob's outcome, and we need only one copy of the announcement register. In this case, the general form for the Kraus operators from Eq. (7) now becomes

$$K_\alpha = \sum_x |r(\alpha, \alpha, x)\rangle_Z \otimes \sqrt{\sum_y P_{(\alpha,x)}^A \otimes P_{(\alpha,y)}^B} \otimes |\alpha\rangle_{\tilde{A}}, \quad (\text{A3})$$

while Eq. (8) becomes

$$K'_{\alpha,w} = \sum_x |r(\alpha, \alpha, x)\rangle_Z \otimes \sqrt{\sum_{x \oplus y = w} P_{(\alpha,x)}^A \otimes P_{(\alpha,y)}^B} \otimes |\alpha\rangle_{\tilde{A}} \otimes |w\rangle_W, \quad (\text{A4})$$

where  $\alpha, \beta$  denotes basis choice, and  $x, y$  denotes measurement outcomes. Alice and Bob's POVMs are given by  $P^A = \{P_{(\alpha,x)}^A\}$  and  $P^B = \{P_{(\alpha,y)}^B\}$ . Since Alice and Bob throw away all signals that have basis mismatch, the set of operators generating the  $\mathcal{G}$  map can be given by  $\{K_\alpha\}$ , and the set of operators generating  $\mathcal{G}'$  is given by  $\{K'_{\alpha,w}\}$ . The  $\mathcal{Z}$  map has Kraus operators  $\{Z_i\}$  given by  $Z_i = |i\rangle\langle i|_Z \otimes I_{AB\tilde{A}}$ . Therefore, the final Kraus operators for  $F$  are given by

$$K_Z = \left[ \begin{pmatrix} 1 \\ 0 \end{pmatrix}_Z \otimes \sqrt{p_z} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_Z \otimes \sqrt{p_z} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_A \right] \otimes \sqrt{p_z} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_B \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\tilde{A}},$$

$$K_X = \left[ \begin{pmatrix} 1 \\ 0 \end{pmatrix}_Z \otimes \sqrt{\frac{p_x}{2}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_Z \otimes \sqrt{\frac{p_x}{2}} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}_A \right] \otimes \sqrt{p_x} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_B \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\tilde{A}}, \quad (\text{A5})$$

and

$$Z_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \mathbb{I}_{\dim_A \times \dim_B \times 2},$$

$$Z_2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \otimes \mathbb{I}_{\dim_A \times \dim_B \times 2}. \quad (\text{A6})$$

The operators for  $F'$  can be constructed from Eq. (A4).

## 2. WCP decoy BB84

Along with source replacement, we use the squashing model from Ref. [41] to squash Bob's system to three dimensions. Since we generate key only

from the single-photon pulses, Alice's POVMs are given by  $\{P_{(Z,0)} = p_z |0\rangle\langle 0|, P_{(Z,1)} = p_z |1\rangle\langle 1|, P_{(X,0)} = p_x |+\rangle\langle +|, P_{(X,1)} = p_x |-\rangle\langle -|\}$ , with  $p_z = p_x = 1/2$ . Bob's POVMs are given by

$$P_{(Z,0)}^B = p_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_{(Z,1)}^B = p_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$P_{(X,0)}^B = \frac{p_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad P_{(X,1)}^B = \frac{p_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix},$$

$$P_{\perp}^B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (\text{A7})$$

with  $p_x = p_z = 1/2$ . Here, the first column corresponds to the vacuum subspace, while the second and third columns make up the qubit subspace. Again, from the discussion in Ref. [42], we can remove certain registers created by the generic form of the Kraus operators in Eq. (7). After removing these registers, the form of the Kraus operators is given by Eq. (A3). Therefore, the final Kraus operators for  $F$  are given by

$$K_Z = \left[ \begin{pmatrix} 1 \\ 0 \end{pmatrix}_Z \otimes \sqrt{p_z} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_Z \otimes \sqrt{p_z} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_A \right] \otimes \sqrt{p_z} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_B \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\tilde{A}},$$

$$K_X = \left[ \begin{pmatrix} 1 \\ 0 \end{pmatrix}_Z \otimes \sqrt{\frac{p_x}{2}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_Z \otimes \sqrt{\frac{p_x}{2}} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}_A \right] \otimes \sqrt{p_x} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_B \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\tilde{A}}, \quad (\text{A8})$$

and

$$Z_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \mathbb{I}_{\dim_A \times \dim_B \times 2},$$

$$Z_2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \otimes \mathbb{I}_{\dim_A \times \dim_B \times 2}. \quad (\text{A9})$$

The operators for  $F'$  can be constructed from Eq. (A4).

## APPENDIX B: BELL-DIAGONAL STATES

For Bell-diagonal states, we can show that the announcement of the location of errors  $W$  leaks no new

information to Eve, by showing that Eve's state is block-diagonal in  $W$  anyway. In the Bell-diagonal case, the state shared between Alice and Bob can be written as

$$\begin{aligned} \rho_{AB} = & \lambda_0 |\phi_+\rangle \langle \phi_+| + \lambda_1 |\phi_-\rangle \langle \phi_-| \\ & + \lambda_2 |\psi_+\rangle \langle \psi_+| + \lambda_3 |\psi_-\rangle \langle \psi_-|, \end{aligned} \quad (\text{B1})$$

where  $|\phi_{+/-}\rangle$  and  $|\psi_{+/-}\rangle$  are the Bell states, and the  $\lambda_i$ s are related to the quantum bit error rate via  $Q_Z = \lambda_3 + \lambda_4$ ,  $Q_X = \lambda_2 + \lambda_4$ , and  $Q_Y = \lambda_2 + \lambda_3$ . We can assume Eve holds a purification of the form

$$\begin{aligned} |\psi\rangle_{ABE} = & \sqrt{\lambda_0} |\phi_+\rangle |e_0\rangle + \sqrt{\lambda_1} |\phi_-\rangle |e_1\rangle \\ & + \sqrt{\lambda_2} |\psi_+\rangle |e_2\rangle + \sqrt{\lambda_3} |\psi_-\rangle |e_3\rangle, \end{aligned} \quad (\text{B2})$$

where  $|e_i\rangle$  are orthonormal basis vectors for Eve's system. Let us suppose Alice and Bob measure in the basis  $\alpha \in \{X, Z\}$ . The (unnormalized) state after the measurement is given by

$$\rho_{XYE}^{(\alpha)} = \sum_{x,y \in \{0,1\}} |x\rangle \langle x| \otimes |y\rangle \langle y| \otimes \rho_E^{(\alpha),x,y}, \quad (\text{B3})$$

where  $\rho_E^{(\alpha),x,y} = \text{Tr}[(P_{(\alpha,x)}^A \otimes P_{(\alpha,y)}^B \otimes I_E) |\psi\rangle \langle \psi|_{ABE}]$ . A simple calculation shows that the support of  $(\rho_E^{(\alpha),0,0}, \rho_E^{(\alpha),1,1})$  is orthogonal to the support of  $(\rho_E^{(\alpha),1,0}, \rho_E^{(\alpha),0,1})$ . Thus, we can conclude that Eve can be assumed to always know the value of  $x \oplus y$  for the entire raw key, if the state shared between Alice and Bob is Bell-diagonal. In fact, the above discussion is also true when Alice and Bob measure in the  $Y$  basis, and is therefore also applicable to the six-state protocol.

### APPENDIX C: TWIRLING REDUCES KEY RATE

For our protocol,  $f(\rho) = S(Z | E\tilde{A}\tilde{B})_\rho$ . One can always expand

$$\begin{aligned} S(Z | E\tilde{A}\tilde{B}) &= \sum_{\alpha,\beta} \text{Prob}(\alpha, \beta) S(Z | E, \tilde{A} = \alpha, \tilde{B} = \beta) \\ &= \sum_{\alpha} \text{Prob}(\alpha) S(Z | E, \tilde{A} = \alpha), \end{aligned} \quad (\text{C1})$$

where we have used the fact that  $Z = \perp$  for basis mismatch, and those signals are thrown away. Now, let

$$\mathcal{T}(\rho) = \frac{1}{4} \sum_{i=1}^4 \rho_i = \frac{1}{4} \sum_{i=1}^4 (\sigma_i \otimes \sigma_i) \rho (\sigma_i \otimes \sigma_i)^\dagger. \quad (\text{C2})$$

Then,

$$\begin{aligned} S(Z | E, \tilde{A} = \alpha)_{\mathcal{T}(\rho)} &\leq \frac{1}{4} \sum_{i=1}^4 S(Z | E, \tilde{A} = \alpha)_{\rho_i} \\ &= \frac{1}{4} \sum_{i=1}^4 S(Z | E, \tilde{A} = \alpha)_\rho \\ &= S(Z | E, \tilde{A} = \alpha)_\rho, \end{aligned} \quad (\text{C3})$$

where we have used the linearity of  $\mathcal{T}$  and the concavity of conditional entropy in the first inequality. The second line follows from the fact that the action of the Pauli operators on  $\rho$  either leave the measurements performed ( $X, Y$ , or  $Z$ ) unchanged, or flip the outcomes, neither of which can affect the entropy. Combining Eqs. (C1) and (C3), we obtain Eq. (11), which is required for the reduction to Bell-diagonal states.

### APPENDIX D: DECOY ANALYSIS

The decoy analysis in this work is similar to that from Ref. [26], with small changes in notation, and is included here for the sake of completeness. For a phase-randomized weak coherent pulse, the state is diagonal in photon number and follows the Poisson probability distribution

$$p_{\mu_i}(n) = \frac{\mu_i^n}{n!} e^{-\mu_i}. \quad (\text{D1})$$

For any statistic  $\gamma_{y|x}$ , one can then write

$$\gamma_{y|x}^{\mu_i} = \sum_n p_{\mu_i}(n) \gamma_{y|x}^n, \quad (\text{D2})$$

where  $\gamma_{y|x}^{\mu_i}$  denotes the probability of Bob obtaining outcome  $y$  given that Alice sent signal  $x$  and intensity  $\mu_i$ . If one uses multiple intensities, then one can use the following set of equations,

$$\begin{aligned} \gamma_{y|x}^{\mu_i} &\leq \sum_{n \leq N} p_{\mu_i}(n) \gamma_{y|x}^n + \left(1 - \sum_{n \leq N} p_{\mu_i}(n)\right), \\ \gamma_{y|x}^{\mu_i} &\geq \sum_{n \leq N} p_{\mu_i}(n) \gamma_{y|x}^n, \end{aligned} \quad (\text{D3})$$

to obtain upper and lower bounds on  $\gamma_{y|x}^1$ :

$$\gamma_{y|x}^{1,L} \leq \gamma_{y|x}^1 \leq \gamma_{y|x}^{1,U}, \quad \forall x, y. \quad (\text{D4})$$

Noting that we can now compute bounds on  $\gamma_{x,y}^1 = \text{Pr}(x) \gamma_{y|x}^1$ , we obtain bounds on all the single-photon statistics for any particular coarse-graining, which we refer

to as

$$S'_1 = \{\rho \in H_+ | \gamma_k^{1,L} \leq \text{Tr}(\Gamma_k \rho) \leq \gamma_k^{1,U}, \forall k\}, \quad (\text{D5})$$

where  $\gamma_k^1$  means the  $k$ th statistics obtained from one-photon signals, and the range of  $k$  depends on the exact nature of the coarse-graining. **Objective function.** The state shared between Alice and Bob after source replacement can be assumed to be block-diagonal in the photon number of Alice's signal, given by

$$\rho_{AA_S B} = \sum_n p_n |n\rangle \langle n|_{A_S} \otimes \rho_{AB}^{(n)}, \quad (\text{D6})$$

where  $A$  and  $B$  are Alice and Bob's systems, and  $A_S$  is a shield system. In such cases, the objective function can be shown to satisfy [43]

$$\min_{\rho \in \mathcal{S}} f(\rho) = \sum_n p_n \min_{\rho_{AB}^{(n)} \in \mathcal{S}'_n} f(\rho_{AB}^{(n)}). \quad (\text{D7})$$

For polarization-encoded phase-randomized pulses, Eve can perform a photon-number-splitting attack [44]. This implies that no key can be generated for  $n > 1$  in the above equation. Therefore, we have

$$\begin{aligned} \min_{\rho \in \mathcal{S}} f(\rho) &= p_0 \min_{\rho_{AB}^{(0)} \in \mathcal{S}'_0} f(\rho_{AB}^{(0)}) + p_1 \min_{\rho_{AB}^{(1)} \in \mathcal{S}'_1} f(\rho_{AB}^{(1)}) \\ &\geq p_1 \min_{\rho_{AB}^{(1)} \in \mathcal{S}'_1} f(\rho_{AB}^{(1)}). \end{aligned} \quad (\text{D8})$$

In this work, we will use the second expression above. Since one does not know the exact single-photon statistics, but rather knows bounds on them due to decoy analysis, the optimization problem is then given by

$$F = \min_{\rho \in \mathcal{S}'_1(G)} f(\rho), \quad (\text{D9})$$

$$S'_1 = \{\rho \in H_+ | \gamma_k^{1,L} \leq \text{Tr}(\Gamma_k \rho) \leq \gamma_k^{1,U}, \forall k\}.$$

- [1] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, *Theor. Comput. Sci.* **560**, 7 (2014).
- [2] D. Bruss, Optimal Eavesdropping in Quantum Cryptography with Six States, *Phys. Rev. Lett.* **81**, 3018 (1998).
- [3] C. H. Bennett, G. Brassard, and N. D. Mermin, Quantum Cryptography Without Bell's Theorem, *Phys. Rev. Lett.* **68**, 557 (1992).
- [4] G. Brassard and L. Salvail, Secret-key reconciliation by public discussion, *Lecture Notes in Computer Science LNCS* **765**, 410 (1994).
- [5] A. Reis, Quantum Key Distribution Post Processing - A study on the Information Reconciliation Cascade Protocol, (2019).

- [6] T. B. Pedersen and M. Toyran, High performance information reconciliation for QKD with CASCADE, *Quantum Inf. Comput.* **15**, 419 (2015).
- [7] J. Martinez-Mateo, C. Pacher, M. Peev, A. Ciurana, and V. Martin, Demystifying the information reconciliation protocol Cascade, *Quantum Inf. Comput.* **15**, 453 (2015).
- [8] D. Elkouss, A. Leverrier, R. Alleaume, and J. J. Boutros, in *2009 IEEE International Symposium on Information Theory, Seoul, Korea (South)* (2009), p. 1879.
- [9] T. Calver, M. Grimaila, and J. Humphries, in *Proceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research (CSIRW'11)* (Association for Computing Machinery, New York, NY, USA, 2011).
- [10] H.-K. Mao, Q. Li, and P.-L. Hao, High performance reconciliation for practical quantum key distribution systems, *arxiv:2101.12565v2* (2021).
- [11] C. Pacher, P. Grabenweger, J. Martinez-Mateo, and V. Martin, in *2015 IEEE International Symposium on Information Theory (ISIT)* (2015), p. 730.
- [12] C. Erven, C. Couteau, R. Laflamme, and G. Weihs, Entangled quantum key distribution over two free-space optical links, *Opt. Express* **16**, 16840 (2008).
- [13] A. R. Dixon, J. F. Dynes, M. Lucamarini, B. Fröhlich, A. W. Sharpe, A. Plews, W. Tam, Z. L. Yuan, Y. Tanizawa, H. Sato, S. Kawamura, M. Fujiwara, M. Sasaki, and A. J. Shields, Quantum key distribution with hacking countermeasures and long term field trial, *Sci. Rep.* **7**, 1978 (2017).
- [14] X. Su, W. Wang, Y. Wang, X. Jia, C. Xie, and K. Peng, Continuous variable quantum key distribution based on optical entangled states without signal modulation, *EPL (Europhysics Letters)* **87**, 20005 (2009).
- [15] C. Gobby, Z. L. Yuan, and A. J. Shields, Quantum key distribution over 122 km of standard telecom fiber, *Appl. Phys. Lett.* **84**, 3762 (2004).
- [16] T. B. H. Tentrup, W. M. Luiten, R. van der Meer, P. Hooijschuur, and P. W. H. Pinkse, Large-alphabet quantum key distribution using spatially encoded light, *New J. Phys.* **21**, 123044 (2019).
- [17] S. Lorenz, N. Korolkova, and G. Leuchs, Continuous-variable quantum key distribution using polarization encoding and post selection, *Appl. Phys. B* **79**, 273 (2004).
- [18] A. Winick, N. Lütkenhaus, and P. J. Coles, Reliable numerical key rates for quantum key distribution, *Quantum* **2**, 77 (2018).
- [19] R. Renner, Security of quantum key distribution, *IEEE Access* **4**, 724 (2016).
- [20] M. Christandl, R. König, and R. Renner, Postselection Technique for Quantum Channels with Applications to Quantum Cryptography, *Phys. Rev. Lett.* **102**, 1 (2009).
- [21] I. George, J. Lin, and N. Lütkenhaus, Numerical calculations of the finite key rate for general quantum key distribution protocols, *Phys. Rev. Res.* **3** (2021).
- [22] V. Scarani and R. Renner, Quantum Cryptography with Finite Resources: Unconditional Security Bound for Discrete-Variable Protocols with One-Way Postprocessing, *Phys. Rev. Lett.* **100**, 200501 (2008).
- [23] V. Scarani and R. Renner, in *Theory of Quantum Computation, Communication, and Cryptography*, edited by

- Y. Kawano and M. Mosca (Springer, Berlin, Heidelberg, 2008), p. 83.
- [24] M. Tomamichel and A. Leverrier, A largely self-contained and complete security proof for quantum key distribution, *Quantum* **1**, 14 (2017).
- [25] M. Curty, M. Lewenstein, and N. Lütkenhaus, Entanglement as a Precondition for Secure Quantum Key Distribution, *Phys. Rev. Lett.* **92**, 217903 (2004).
- [26] W. Wang and N. Lütkenhaus, Numerical security proof for the decoy-state BB84 protocol and measurement-device-independent quantum key distribution resistant against large basis misalignment, *Phys. Rev. Res.* **4**, 043097 (2022).
- [27] I. Devetak and A. Winter, Distillation of secret key and entanglement from quantum states, *Proc. R. Soc. A: Math., Phys. Eng. Sci.* **461**, 207 (2005).
- [28] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, Secure quantum key distribution with realistic devices, *Rev. Mod. Phys.* **92**, 025002 (2020).
- [29] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, The security of practical quantum key distribution, *Rev. Mod. Phys.* **81**, 1301 (2009).
- [30] A. Ferenczi and N. Lütkenhaus, Symmetries in quantum key distribution and the connection between optimal attacks and optimal cloning, *Phys. Rev. A* **85**, 052310 (2012).
- [31] S. Watanabe, R. Matsumoto, T. Uyematsu, and Y. Kawano, Key rate of quantum key distribution with hashed two-way classical communication, *Phys. Rev. A - At., Mol. Opt. Phys.* **76** (2007).
- [32] D. Tupkary, Master's thesis, University of Waterloo, 2022.
- [33] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824 (1996).
- [34] C. C. W. Lim, M. Curty, N. Walenta, F. Xu, and H. Zbinden, Concise security bounds for practical decoy-state quantum key distribution, *Phys. Rev. A* **89**, 022307 (2014).
- [35] H.-K. Lo, X. Ma, and K. Chen, Decoy State Quantum Key Distribution, *Phys. Rev. Lett.* **94**, 230504 (2005).
- [36] D. Rusca, A. Boaron, F. Grünenfelder, A. Martin, and H. Zbinden, Finite-key analysis for the 1-decoy state QKD protocol, *Appl. Phys. Lett.* **112**, 171104 (2018).
- [37] P. Rice and J. Harrington, Numerical analysis of decoy state quantum key distribution protocols, [arxiv:0901.0013](https://arxiv.org/abs/0901.0013) [quant-ph] (2009).
- [38] W.-Y. Hwang, Quantum Key Distribution with High Loss: Toward Global Secure Communication, *Phys. Rev. Lett.* **91**, 057901 (2003).
- [39] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo, Practical decoy state for quantum key distribution, *Phys. Rev. A* **72**, 012326 (2005).
- [40] X.-B. Wang, Beating the Photon-Number-Splitting Attack in Practical Quantum Cryptography, *Phys. Rev. Lett.* **94**, 230503 (2005).
- [41] O. Gittsovich, N. J. Beaudry, V. Narasimhachar, R. R. Alvarez, and T. Moroder, Squashing model for detectors and applications to quantum-key-distribution protocols, *Phys. Rev. A* **89**, 012325 (2014).
- [42] J. Lin, T. Upadhyaya, and N. Lütkenhaus, Asymptotic Security Analysis of Discrete-Modulated Continuous-Variable Quantum Key Distribution, *Phys. Rev. X* **9**, 41064 (2019).
- [43] N. K. H. Li and N. Lütkenhaus, Improving key rates of the unbalanced phase-encoded BB84 protocol using the flag-state squashing model, *Phys. Rev. Res.* **2**, 043172 (2020).
- [44] N. Lütkenhaus and M. Jähma, Quantum key distribution with realistic states: Photon-number statistics in the photon-number splitting attack, *New J. Phys.* **4** (2002).