Dimension Reduction in Quantum Key Distribution for Continuous- and Discrete-Variable Protocols

Twesh Upadhyaya[1,*] Thomas van Himbeeck,[1,2] Jie Lin[1] and Norbert Lütkenhaus[1]

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

[2]*Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario M5S 3G4, Canada*

We develop a method to connect the infinite-dimensional description of optical continuous-variable quantum key distribution (QKD) protocols to a finite-dimensional formulation. The secure key rates of the optical QKD protocols can then be evaluated using recently developed reliable numerical methods for key-rate calculations. We apply this method to obtain asymptotic key rates for discrete-modulated continuous-variable QKD protocols, which are of practical significance due to their experimental simplicity and potential for large-scale deployment in quantum-secured networks. Importantly, our security proof does not require the photon-number cutoff assumption relied upon in previous works. We also demonstrate that our method can provide practical advantages over the flag-state squasher when applied to discrete-variable protocols.

I. INTRODUCTION

Quantum key distribution (QKD) [1,2] enables two remote parties, Alice and Bob, to establish information theoretically secure keys even in the presence of an eavesdropper, Eve. These keys can then be used in many other cryptographic applications, such as the one-time pad. To prove QKD to be secure with a specified key rate, we need to assume a quantum-mechanical model for Alice and Bob's devices, but do not have to assume anything about the processing power available to the eavesdropper [3]. Reviews of QKD protocols can be found in Refs. [4–6].

For a given QKD protocol, the goal of a security proof is to find a lower bound on the secure key rate. Analytical methods for this task can be very involved, tend to be restricted to symmetric protocols, and can introduce looseness in the lower bounds. These issues are ameliorated by the recent development of tight, reliable numerical methods for finding secure key rates [7,8]. At a high level, these methods determine the key rate by solving a particular convex optimization over the set of quantum states that could be held by Alice and Bob. Thus, when the bipartite Hilbert space is infinite dimensional, the numerical

methods cannot be used directly. Fortunately, for many discrete-variable (DV) protocols, the numerical methods can be applied by using the squashing map [9–11] or the more general flag-state squasher [12] to reduce the problem to finite dimensions. However, these squashing approaches do not seem applicable to continuous-variable (CV) protocols. Additionally, even for DV protocols, the flag-state squasher can have challenging runtimes [13].

Discrete-modulated continuous-variable QKD (DMCV QKD) is a family of protocols that utilize existing telecommunication infrastructure, including homodyne or conjugate homodyne detection [14–16]. They are thus promising candidates for deployment in large-scale quantum-secured networks. In comparison to Gaussian-modulated CVQKD [17–19], discrete modulation is less demanding on the source modulator and on the error-correction protocols, yet is expected to achieve similar key rates. It is thus of interest to establish security proofs for DMCVQKD. Of particular interest is DMCVQKD with four or more modulated states, which is expected to outperform protocols with a smaller constellation. While there are analytic asymptotic security proofs of DMCVQKD with two [20] or three [21] modulated states, they are difficult to generalize to more states. Recent works have numerically studied asymptotic security proofs for DMCVQKD with any number of modulated states [22,23]. However, these approaches assume the state is finite dimensional, known as the photon-number cutoff assumption. Thus, while these results seem numerically plausible, they do not constitute a rigorous asymptotic security proof, as the cutoff

---

*twesh.upadhyaya@uwaterloo.ca

assumption cannot be justified. A full finite-key analysis of binary-modulated DMCVQKD has also been recently completed in Ref. [24].

The main contribution of this paper is a method to tightly lower bound the key rate of an infinite-dimensional QKD protocol in terms of a finite-dimensional convex optimization. In combination with existing numerical tools for solving finite-dimensional convex optimizations, this enables us to find tight, reliable key rates for general device-dependent QKD protocols in infinite-dimensional Hilbert spaces. Our dimension reduction method can also be applied to study other tasks in quantum information, such as entanglement verification [25].

As a result, our method can provide a complete asymptotic security proof for discrete-modulated continuous-variable protocols with any number of modulated states, with tight key rates and without relying on the photon-number cutoff assumption. While our focus in this work is on calculating asymptotic key rates, we expect key elements of our method to lift to a finite-key analysis.

Our dimension reduction method also provides an alternative approach to study protocols admitting a flag-state squasher. We consider unbalanced phase-encoded BB84 as an example, and show that our method can have an improved runtime compared to the flag-state squasher, while providing similar results.

The remainder of this paper is structured as follows. In Sec. II, we review the basic steps of a QKD protocol and how the key rate can be formulated as a convex optimization. In Sec. III, we develop our framework for dimension reduction, in more generality than is needed for QKD. In Sec. IV, we then specialize our general method to asymptotic key rate calculations. In Sec. V, we show how to implement the relevant optimizations numerically. In Sec. VI, we calculate key rates for DMCVQKD, including modeling postselection and trusted noise. In Sec. VII we compare our method to the flag-state squasher. Finally, we provide concluding remarks and avenues for future work in Sec. VIII. Certain technical details are relegated to the appendices.

II. BACKGROUND: QKD PROTOCOLS AND SECURITY ANALYSIS

A. Generic QKD protocol steps

We first review the basic steps of a generic QKD protocol. Alice and Bob have access to an uncharacterized quantum channel and an authenticated, public classical channel. By the source-replacement scheme [26–29], any prepare-and-measure (P&M) protocol can be equivalently viewed as an entanglement-based (EB) one. Thus, without loss of generality we consider entanglement-based protocols.

1. Alice and Bob establish a quantum state ρ_{AB}.

2. Alice and Bob measure their subsystems with positive operator-valued measures (POVMs) $\{P_A^i\}$ and $\{P_B^j\}$. To each outcome i, j, they associate two pieces of classical data: a public announcement a_i, b_j and a private measurement result α_i, β_j. The respective alphabets from which the values are drawn are $S_a, S_b, S_\alpha, S_\beta$. We can think of the classical announcements as partitioning the data.

After repeating the previous two steps for a large number of rounds, Alice and Bob proceed to the classical phase.

3. Alice and Bob choose a random subset of the rounds to use for parameter estimation. For these testing rounds, they announce their public and private results. This allows them to determine the expectations γ_i of some testing observables Γ_i.

4. Alice and Bob announce their public data. Based on the joint announcements, they may decide to discard some rounds. This is represented by a binary function $d : S_a \times S_b \to \{0, 1\}$; where $d = 0$ if the signal is kept and $d = 1$ if discarded.

5. Based on the public announcements and their private data, one party performs the *key map*. When Alice (Bob) performs the key map, it is conventionally referred to as direct (reverse) reconciliation. In the following discussion, we consider the case where Bob performs the key map. For a k-ary key, the key map is a function $g : S_a \times S_b \times S_\beta \to \{0, 1, \ldots, k-1, \perp\}$. The $\perp$ symbol is only used to flag the discarded or *sifted* signals, so $g(a, b, \beta) = \perp \iff d(a, b) = 1$.

6. Alice and Bob then perform error correction over the classical channel to get Alice's data to agree with the sifted key established by Bob.

7. Alice and Bob perform privacy amplification using a two-universal hash function to obtain the final shared secret key.

In practice, the discarded signals are simply removed before performing privacy amplification. We include them with the discard flag $\perp$ only to formulate the protocol in a trace-preserving manner.

This description of an EB protocol is general. When modeling the EB version of a P&M protocol, some steps can be made more specific. Suppose Alice prepares signal states $|\psi_i\rangle$ with probability $p(i)$. In the source-replacement scheme, this is modeled as Alice preparing the state $\tau_{AA'} = \sum_{ij} \sqrt{p(i)p(j)} |i\rangle \langle j|_A \otimes |\psi_i\rangle \langle \psi_j|_{A'}$, sending system A' to Bob, and measuring with the POVM $\{|i\rangle \langle i|_A\}$. As Eve cannot access Alice's lab, the reduced density matrix is known. Thus, for P&M protocols there is an additional constraint in parameter estimation, namely $\rho_A = \tau_A = \sum_{ij} \sqrt{p(i)p(j)} \langle \psi_j | \psi_i \rangle |i\rangle \langle j|$. Note that τ_A is closely related to the Gram matrix of signal states.

B. Postprocessing channel

In order to evaluate the key rate, we formally define the postprocessing steps as a quantum channel. We introduce registers $\tilde{A}$ and $\tilde{B}$, which hold the public announcements, $\overline{A}$ and $\overline{B}$, which hold the private measurement data, and Z, which holds the result of the key map. Without loss of generality, Eve has access to the register E purifying ρ_{AB} and the public information.

Alice and Bob's measurement can be described by a channel Φ_M^{AB} that is simply given by $\mathrm{Tr}_{AB}(\Phi_T^{AB})$, where

$$\Phi_T^{AB}(\rho_{ABE}) = \sum_{i,j} |a_i\rangle \langle a_i|_{\tilde{A}} \otimes |\alpha_i\rangle \langle \alpha_i|_{\overline{A}}$$
$$\otimes |b_j\rangle\langle b_j|_{\tilde{B}} \otimes |\beta_j\rangle\langle \beta_j|_{\overline{B}}$$
$$\otimes \left[\left(\sqrt{P_A^i} \otimes \sqrt{P_B^j}\right) \rho_{ABE} \left(\sqrt{P_A^i} \otimes \sqrt{P_B^j}\right) \right]. \tag{1}$$

The action of the key map can be represented by the isometry

$$V = \sum_{S_a, S_b, S_\beta} |g(a, b, \beta)\rangle_Z \otimes |a\rangle \langle a|_{\tilde{A}} \otimes |b\rangle \langle b|_{\tilde{B}} \otimes |\beta\rangle \langle \beta|_{\overline{B}}. \tag{2}$$

The final state between all parties is then $V\Phi_M^{AB} (\rho_{ABE}) V^\dagger$. We can pull the partial trace in the measurement channel Φ_M^{AB} through the isometry V. Then, the final classical-quantum state between the register holding the result of the key map and Eve is

$$\sigma_{Z[E]} = \mathrm{Tr}_{A\overline{A}B\overline{B}}[V\Phi_T^{AB}(\rho_{ABE})V^\dagger], \tag{3}$$
$$\equiv \Phi(\rho_{ABE}). \tag{4}$$

The channel Φ characterizes the postprocessing steps and $[E]$ denotes the composite register $E\tilde{A}\tilde{B}$. Note that discarded signals will not contribute to the key rate, as $\sigma_{Z[E]}$ is block diagonal in the classical announcements. To compute Φ, it is thus simpler and equivalent to not apply the POVM elements leading to discarded outcomes, rather than use a discard symbol [23]. In this case, the postprocessing map is completely positive and trace nonincreasing.

C. Asymptotic key-rate formula

The secure key rate of a QKD protocol, in the asymptotic limit and assuming independent and identically distributed (IID) rounds (collective attack), is given by the Devetak-Winter formula [30]. One can then aim to lift this to the key rate under coherent attacks, for example, using a quantum de Finetti theorem [31], and ultimately to a fully composable security proof incorporating finite-size effects.

Under collective attacks, for a given ρ_{AB}, the asymptotic key rate is

$$R^\infty = I(Z : X) - \chi(Z|[E]), \tag{5}$$

where I denotes the mutual information and χ denotes the Holevo information. The quantities are evaluated on the postprocessed state after a single round, and X refers to the party who does not perform the key map. In general, the state ρ_{AB} is unknown. However, it is constrained by Alice and Bob through testing. The Devetak-Winter formula should be evaluated on the worst-case state compatible with these constraints.

Determining the key rate of a protocol can then be reformulated as a convex optimization problem [7]. The Devetak-Winter formula can be rearranged as

$$R^\infty = H(Z|[E]) - H(Z) + I(Z : X), \tag{6}$$
$$= H(Z|[E]) - H(Z|X). \tag{7}$$

Only the first term needs to be optimized over. The second term is replaced by the actual error-correction cost δ_{EC}^{leak}, which is the number of bits leaked per round. For realistic error correction, δ_{EC}^{leak} will be larger than the Shannon limit $H(Z|X)$. The key rate is thus

$$R^\infty = \min_{\rho_{AB} \in \mathbf{S}^{QKD}} [f^{QKD}(\rho_{AB})] - \delta_{EC}^{leak}, \tag{8}$$

where the objective function is the conditional entropy evaluated on the purified state after postprocessing

$$f^{QKD}(\rho_{AB}) = H(Z|[E])_{\Phi(\rho_{ABE})}, \tag{9}$$

and the set $\mathbf{S}^{QKD}$ is defined by the parameter estimation Alice and Bob perform, as well as the reduced density matrix constraint for P&M protocols [8]. That is,

$$\mathbf{S}^{QKD} = \{\rho \in \mathrm{Pos}\,(\mathcal{H}_{AB}) : \mathrm{Tr}(\rho) = 1,$$
$$\mathrm{Tr}_B(\rho) = \tau_A, \tag{10}$$
$$\mathrm{Tr}(\rho\Gamma_i) = \gamma_i\}.$$

As both the objective function and feasible set are convex [8], Eq. (8) is a convex optimization.

When the Hilbert space $\mathcal{H}_{AB}$ is finite dimensional, this problem can be reliably solved numerically [7,8]. However, when the Hilbert space is infinite dimensional, it is clearly not possible to directly solve this optimization numerically. We develop a general method that, for all QKD protocols, allows us to compute tight lower bounds on the asymptotic key rate by relating the infinite-dimensional optimization to a finite-dimensional one. We can then numerically solve the finite-dimensional problem using the methods of Ref. [8] to get tight lower bounds on the key rate for protocols where the state lives in an infinite-dimensional Hilbert space.

III. GENERAL FRAMEWORK

In this section, we develop a general method to lower bound infinite-dimensional optimizations of the sort seen in the previous section. In fact, we present our results in more generality than used for finding QKD key rates. Our general method may be applied to other scenarios in quantum information, such as entanglement verification [25].

Our notational conventions are as follows. For a Hilbert space $\mathcal{H}$, let $D(\mathcal{H})$ ($\tilde{D}(\mathcal{H})$) be the set of (sub)normalized density operators on $\mathcal{H}$. That is, $\tilde{D}(\mathcal{H}) = \text{Pos}(\mathcal{H}) \cap \mathcal{T}_1$, where $\text{Pos}(\mathcal{H})$ is the set of positive semidefinite operators on $\mathcal{H}$, and $\mathcal{T}_1$ is the set of trace-class operators with trace no greater than 1. We use tildes to denote operators that are subnormalized.

A. Problem setup and definitions

We begin with some definitions. Let $\mathcal{H}_\infty$ be a separable Hilbert space, which may be infinite dimensional. Let $\mathbf{S}_\infty$ be a convex subset of $\tilde{D}(\mathcal{H}_\infty)$. Finally, let f be a convex function from $\tilde{D}(\mathcal{H}_\infty)$ to $\mathbb{R}$. The infinite-dimensional optimization we consider is

$$\min_{\tilde{\rho} \in \mathbf{S}_\infty} f(\tilde{\rho}). \tag{11}$$

We assume that there exists a feasible operator $\tilde{\rho}^\infty$ achieving the optimum. This assumption holds in the practical case where the objective function f is continuous and $\mathbf{S}_\infty$ is compact. Our goal then is to find a lower bound on $f(\tilde{\rho}^\infty)$. Our strategy to do this is to relate the infinite-dimensional optimization to a suitably chosen finite-dimensional optimization, which can then be solved numerically. Note that in the process we choose certain mathematical objects freely; how to choose them effectively is the focus of the following sections.

Choose $\mathcal{H}_N$ to be any finite-dimensional subspace of $\mathcal{H}_\infty$. Π is defined as the projector onto this subspace, and $\bar{\Pi} \equiv \mathbb{1} - \Pi$. Choose $\mathbf{S}_N$ to be a nonempty convex subset of $\tilde{D}(\mathcal{H}_N)$ satisfying

$$\Pi \mathbf{S}_\infty \Pi \subseteq \mathbf{S}_N. \tag{12}$$

It is always possible to choose such a set $\mathbf{S}_N$, as $\Pi \tilde{D}(\mathcal{H}_\infty) \Pi \subseteq \tilde{D}(\mathcal{H}_N)$. We can now define the finite-dimensional optimization

$$\min_{\tilde{\rho} \in \mathbf{S}_N} f(\tilde{\rho}). \tag{13}$$

We assume there exists a feasible operator $\tilde{\rho}^N$ achieving this optimum. Again, this assumption holds in the practical case when f is continuous and $\mathbf{S}_N$ is compact.

B. Main theorem

At a high level, our proof method is illustrated in Fig. 1. In order to relate the objective function f evaluated at the
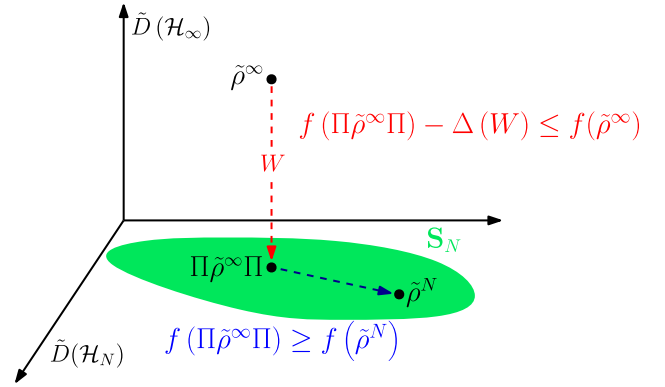

FIG. 1. Pictorial representation of Theorem 1. The set $\mathbf{S}_N$ is chosen to contain the projection $\Pi \tilde{\rho}^\infty \Pi$. This auxiliary variable is used to relate $f(\tilde{\rho}^\infty)$ and $f(\tilde{\rho}^N)$.

two optimal operators, $\tilde{\rho}^\infty$ and $\tilde{\rho}^N$, we introduce an auxiliary variable, $\tilde{\rho}^\Pi \equiv \Pi \tilde{\rho}^\infty \Pi$. This variable is the projection of the infinite-dimensional optimum $\tilde{\rho}^\infty$ onto the chosen finite subspace. We relate the optima to this auxiliary variable separately, and then to each other. For positive operators P and Q, our convention for the fidelity function is $F(P, Q) = \text{Tr}(\sqrt{\sqrt{Q}P\sqrt{Q}})$.

Inequality 1, finite set: We first relate $f(\tilde{\rho}^N)$ and $f(\tilde{\rho}^\Pi)$. By definition, $\tilde{\rho}^\Pi \in \Pi \mathbf{S}_\infty \Pi$. By the containment property introduced in Eq. (12), it follows that $\tilde{\rho}^\Pi \in \mathbf{S}_N$. Thus, $\tilde{\rho}^\Pi$ is feasible for the minimization in Eq. (13). Since $\tilde{\rho}^N$ achieves the minimum, it follows that

$$f(\tilde{\rho}^N) \leq f(\tilde{\rho}^\Pi). \tag{14}$$

Inequality 2, projection: We next relate $f(\tilde{\rho}^\infty)$ and $f(\tilde{\rho}^\Pi)$. To do this, we introduce a property of f relating the fidelity of input states to changes in the function. Define f as being *uniformly close to decreasing under projection (UCDUP) on* $\mathbf{S} \subseteq \tilde{D}(\mathcal{H}_\infty)$ *with correction term* Δ if, for every $\tilde{\sigma} \in \mathbf{S}$, it satisfies

$$F(\tilde{\sigma}, \Pi \tilde{\sigma} \Pi) \geq \text{Tr}(\tilde{\sigma}) - W$$
$$\implies f(\Pi \tilde{\sigma} \Pi) - f(\tilde{\sigma}) \leq \Delta(W), \quad \text{(15)}$$

where Δ is a non-negative, increasing function satisfying $\Delta(0) = 0$. This property has some similarities to uniform continuity with respect to trace distance. In fact, it is implied by the latter. However, because being UCDUP is a weaker condition, we may be able to find smaller Δ than implied by a uniform-continuity bound (see Sec. IV C 1). In some cases, it may also be possible to set $\Delta = 0$. This is the case for bounding the QKD objective function when the key map POVM elements commute with Π (see Sec. IV C 2). It is also the case for some entanglement measures [25].

We can re-express the condition on W in a more useful way. By Winter's gentle measurement lemma [32],

which holds with equality when considering projections, the fidelity can be written as

$$F(\tilde{\sigma}, \Pi\tilde{\sigma}\Pi) = \mathrm{Tr}(\tilde{\sigma}\Pi). \quad (16)$$

Rearranging, it follows that

$$F(\tilde{\sigma}, \Pi\tilde{\sigma}\Pi) \geq \mathrm{Tr}(\tilde{\sigma}) - W \iff \mathrm{Tr}(\tilde{\sigma}\bar{\Pi}) \leq W. \quad (17)$$

Thus, when f is UCDUP on $\mathbf{S}_\infty$ and $\mathrm{Tr}(\tilde{\rho}^\infty\bar{\Pi}) \leq W$,

$$f(\tilde{\rho}^\Pi) - \Delta(W) \leq f(\tilde{\rho}^\infty). \quad (18)$$

We now have all the pieces in place to prove our main theorem.

Theorem 1 (Relating finite- and infinite-dimensional optimizations). *If f is UCDUP on $\mathbf{S}_\infty$ with correction term Δ [Eq. (15)] and $\mathrm{Tr}(\tilde{\rho}^\infty\bar{\Pi}) \leq W$, then*

$$f(\tilde{\rho}^N) - \Delta(W) \leq f(\tilde{\rho}^\infty). \quad (19)$$

Proof. This theorem follows from chaining the inequality in Eq. (14), which relates $f(\tilde{\rho}^N)$ and $f(\tilde{\rho}^\Pi)$, with the one in Eq. (18), which relates $f(\tilde{\rho}^\infty)$ and $f(\tilde{\rho}^\Pi)$. ∎

In order to apply this theorem, in addition to choosing $\mathcal{H}_N$ and $\mathbf{S}_N$, we need to determine an expression for the correction term Δ and a suitable value for W (see Sec. IV). For the latter quantity, as $\tilde{\rho}^\infty$ is unknown, we choose the smallest value of W satisfying

$$W \geq \max_{\tilde{\rho}\in\mathbf{S}_\infty} \mathrm{Tr}(\tilde{\rho}\bar{\Pi}). \quad (20)$$

We thus see that the determination of W itself involves an infinite-dimensional optimization. In practice, this optimization tends to be considerably easier to solve than the original one [Eq. (11)]. In particular, when $\mathbf{S}_\infty$ is the feasible set of a semidefinite program (SDP), known as a spectrahedron, then any solution to the dual problem provides an upper bound. For the QKD protocols we study, we can obtain tight values of W by analytically solving the dual or a relaxed version of the primal problem (see Sec. IV B).

Note that W is not only used to determine the correction term $\Delta(W)$, but also to parametrize the set $\mathbf{S}_N$ we choose (see Sec. IV D).

IV. APPLICATION OF FRAMEWORK TO QKD

Having developed a general method to lower-bound convex optimizations, we now apply it to asymptotic QKD key rate optimizations. Our initial infinite-dimensional optimization is as described in Sec. II C, so that $\mathcal{H}_\infty = \mathcal{H}_{AB}, f = f^{\mathrm{QKD}}$, and $\mathbf{S}_\infty = \mathbf{S}^{\mathrm{QKD}}$. Unless otherwise stated,

we focus on the case where $\mathcal{H}_A$ is finite dimensional. In order to apply Theorem 1, we need four inputs: the subspace $\mathcal{H}_N$ onto which Π projects, the bound on weight outside the subspace W, the correction term Δ, and the finite set $\mathbf{S}_N$. The first two of these depend on the specific QKD protocol, so we give only some general remarks on them in this section. The second two are generic for all protocols, so we derive them in detail.

A. Choose subspace $\mathcal{H}_N$

This step is protocol specific, and the choice of finite subspace $\mathcal{H}_N$ can be made freely. However, for a fixed finite dimension, some subspaces will give a better lower bound than others. Intuitively, we want to choose a subspace that contains most of the weight of $\tilde{\rho}^\infty$. Since this state is unknown, a good heuristic is to choose the subspace containing the most weight of the expected state under a representative channel model (see Sec. VI C 1). This is conceptually similar to assuming a particular channel for the purpose of designing a QKD protocol or error-correcting code. Another consideration is to choose the projection to commute with the objective function POVM elements (see Sec. VI C 2) or the constraint operators (see Sec. IV D). The choice of the dimension of the finite subspace is influenced by the run time of the numerical method used to solve the finite-dimensional optimization.

B. Bound weight W outside subspace

This step is also protocol specific. Note that for $\mathbf{S}_\infty$ as given in Eq. (10), the bound on W shown in Eq. (20) is a semidefinite program. The range of approaches to calculate W is fairly wide, so in place of general strategies, we instead survey some examples. For DMCVQKD, we analytically find a solution to the dual SDP (see Sec. VI C 2). For unbalanced phase-encoded BB84, the monotonicity of cross clicks or double clicks with increasing photon number, along with Markov's inequality, is used to bound W [13].

C. Determine correction term Δ

Recall our objective function f is given in Eq. (9). We show that it is UCDUP and determine the correction term Δ as a function of W. We find a general correction term, which does not depend on the details of the postprocessing map [see Eq. (4)] and is thus applicable to all protocols. We also show that when the postprocessing map satisfies a certain property, which holds for some protocols, we can omit the correction term entirely, i.e., $\Delta = 0$.

1. General case

In this case, we do not make any assumptions about the postprocessing channel Φ in Eq. (4), other than the fact

that it is indeed a channel. The crux of the proof is in the following lemma.

Lemma 1. *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two Hilbert spaces, where the dimension of $\mathcal{H}_A$ is $|A|$ while $\mathcal{H}_B$ can be infinite dimensional. Let $\tilde{\rho}_{AB}, \tilde{\sigma}_{AB} \in \tilde{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be two subnormalized, classical-quantum states with $Tr(\tilde{\rho}_{AB}) \geq Tr(\tilde{\sigma}_{AB})$. If $\frac{1}{2} \| \tilde{\rho}_{AB} - \tilde{\sigma}_{AB} \|_1 \leq \epsilon$, then*

$$H(A|B)_{\tilde{\sigma}_{AB}} - H(A|B)_{\tilde{\rho}_{AB}} \leq \epsilon \log_2 |A|$$
$$+ (1 + \epsilon)h\left(\frac{\epsilon}{1 + \epsilon}\right), \tag{21}$$

where $h(x)$ is the binary entropy function.

Proof. This result is a generalization of Lemma 2 in Ref. [33], which was derived for normalized states. The proof of the extension to subnormalized states is given in Appendix A. ∎

With this lemma in hand, we prove that f is UCDUP on $D(\mathcal{H}_\infty)$ in the following theorem. Note this is sufficient as for any QKD protocol, $\mathbf{S}_\infty \subseteq D(\mathcal{H}_\infty)$.

Theorem 2. *The QKD objective function f is UCDUP on $D(\mathcal{H}_\infty)$ with correction term*

$$\Delta(W) = \left(\sqrt{2W - W^2} \log_2 |Z|\right)$$
$$+ \left(1 + \sqrt{2W - W^2}\right) h\left(\frac{\sqrt{2W - W^2}}{1 + \sqrt{2W - W^2}}\right), \tag{22}$$

where $|Z|$ is the dimension of the key map register.

Proof. As per the definition of UCDUP, let $\sigma_{AB} \in D(\mathcal{H}_\infty)$ be a state satisfying $F(\sigma_{AB}, \Pi\sigma_{AB}\Pi) \geq 1 - W$. Let σ_{ABE} and $\tilde{\sigma}_{ABE}^{\Pi}$ be purifications of σ_{AB} and $\Pi\sigma_{AB}\Pi$, respectively, (we consider a purification of an unnormalized state to not change the trace). By Uhlmann's theorem [34], we can choose the purifications to satisfy

$$F(\sigma_{ABE}, \tilde{\sigma}_{ABE}^{\Pi}) = F(\sigma_{AB}, \Pi\sigma_{AB}\Pi) \tag{23}$$
$$\geq 1 - W. \tag{24}$$

By the monotonicity of fidelity under channels, we have

$$F(\tau_{Z[E]}, \tilde{\tau}_{Z[E]}^{\Pi}) \geq 1 - W, \tag{25}$$

where $\tau_{Z[E]} = \Phi(\sigma_{ABE})$ and $\tilde{\tau}_{Z[E]}^{\Pi} = \Phi(\tilde{\sigma}_{ABE}^{\Pi})$. The Fuchs-van de Graaf inequalities [35], which are valid in infinite dimensions and for subnormalized states, relate

fidelity and trace distance as follows: $1 - F(\tilde{\rho}, \tilde{\sigma}) \leq \frac{1}{2} \| \tilde{\rho} - \tilde{\sigma} \|_1 \leq \sqrt{1 - F(\tilde{\rho}, \tilde{\sigma})^2}$. By the second of these,

$$\frac{1}{2} \left\| \tau_{Z[E]} - \tilde{\tau}_{Z[E]}^{\Pi} \right\|_1 \leq \sqrt{2W - W^2}. \tag{26}$$

Since Φ is trace preserving, $\text{Tr}\left(\tau_{Z[E]}\right) \geq \text{Tr}\left(\tilde{\tau}_{Z[E]}^{\Pi}\right)$. Thus, we can apply Lemma 1 to obtain

$$f\left(\Pi\sigma_{AB}\Pi\right) - f\left(\sigma_{AB}\right) = H(Z|[E])_{\tilde{\tau}_{Z[E]}^{\Pi}} - H(Z|[E])_{\tau_{Z[E]}} \tag{27}$$

$$\leq \sqrt{2W - W^2} \log_2 |Z|$$
$$+ (1 + \sqrt{2W - W^2})h\left(\frac{\sqrt{2W - W^2}}{1 + \sqrt{2W - W^2}}\right). \tag{28}$$

This is precisely the condition we require for f to be UCDUP, so we identify the right-hand side as $\Delta(W)$. ∎

For a key map with k outcomes, the dimension of the key map register is k. For the purpose of this argument, the discard symbol $\perp$ does not count towards a key outcome. The reason for this is that instead of $\perp$, one could use any pre-existing key symbol to flag discarded signals. Since the classical-quantum state $\sigma_{Z[E]}$ between the key map register and Eve is block diagonal in the classical announcements [Eq. (4)], Eve could identify the discarded signals from those public announcements alone. This would leave the value of the objective function unchanged. Indeed, the $\perp$ symbol is only used for clarity in our presentation.

2. Special case: Block-diagonal measurements

If the key map POVM elements are block diagonal with respect to Π and $\bar{\Pi}$, the correction term is zero.

Theorem 3. *Let Φ be defined by the POVMs $\{P_A^i\}$ and $\{P_B^j\}$ and a key map isometry V. If all POVM elements are block diagonal, so that $[P_A^i \otimes P_B^j, \Pi] = 0 \; \forall i,j$, then*

$$\Delta(W) = 0. \tag{29}$$

Proof. Recall that Π only acts on the AB subsystem. Then, $\Pi\rho_{ABE}\Pi$ is a purification of $\Pi\rho_{AB}\Pi$. Π commutes with V as they act on different subsystems, and commutes with all elements of the POVMs $\{P_A^i\}$ and $\{P_B^j\}$ by assumption. By the definition of Φ_T^{AB} [Eq. (1)], it then follows that

$$\Pi V \Phi_T^{AB}(\rho_{ABE}) V^\dagger \Pi = V \Phi_T^{AB}(\Pi \rho_{ABE} \Pi) V^\dagger \tag{30}$$

and analogously for $\bar{\Pi}$.

Defining the channel $\Xi_{AB}(\rho) = \Pi\rho\Pi + \bar{\Pi}\rho\bar{\Pi}$, we have

$$f(\rho_{AB}) = H(Z|[E])_{\Phi(\rho_{ABE})} \tag{31}$$

$$= H(Z|[E])_{\mathrm{Tr}_{A\bar{A}B\bar{B}}\left[V\Phi_T^{AB}(\rho_{ABE})V^\dagger\right]} \tag{32}$$

$$= H(Z|[E])_{\mathrm{Tr}_{A\bar{A}B\bar{B}}\left(\Xi_{AB}\left[V\Phi_T^{AB}(\rho_{ABE})V^\dagger\right]\right)} \tag{33}$$

$$= H(Z|[E])_{\mathrm{Tr}_{A\bar{A}B\bar{B}}\left(V\Phi_T^{AB}[\Xi_{AB}(\rho_{ABE})]V^\dagger\right)} \tag{34}$$

$$\geq H(Z|[E])_{\Phi(\Pi\rho_{ABE}\Pi)} + H(Z|[E])_{\Phi(\bar{\Pi}\rho_{ABE}\bar{\Pi})} \tag{35}$$

$$\geq H(Z|[E])_{\Phi(\Pi\rho_{ABE}\Pi)} \tag{36}$$

$$= f(\Pi\rho_{AB}\Pi), \tag{37}$$

where in Eq. (33) we freely introduce a channel as it is traced out, in Eq. (35) use the concavity of conditional entropy, and in Eq. (36) use the nonnegativity of conditional entropy for classical-quantum states [see Eq. (4)]. ∎

D. Choose finite set $\mathbf{S}_N$

Recall that the feasible set $\mathbf{S}_\infty$ for the infinite-dimensional optimization is given in Eq. (10). The form of this set is common to all protocols. We consider one particular way of choosing $\mathbf{S}_N$, which is to individually expand the constraints of $\mathbf{S}_\infty$, using W as a parameter. Note that in order to have the highest key rates, we want $\mathbf{S}_N$ to be as small as possible, while still satisfying the containment in Eq. (12). Assuming $\sigma \in \mathbf{S}_\infty$, we now list constraints that $\tilde{\sigma}^\Pi \equiv \Pi\sigma\Pi$ must satisfy.

We begin with the trace constraint, $\mathrm{Tr}(\sigma) = 1$, which can be easily expanded. By the definition of W in Eq. (20), $1 - W \leq \mathrm{Tr}(\tilde{\sigma}^\Pi) \leq 1$.

We next consider the expectation constraints, $\mathrm{Tr}(\sigma\Gamma_i) = \gamma_i$. We choose to define the loosened constraints as $\gamma_i^{\min} \leq \mathrm{Tr}(\tilde{\sigma}^\Pi\Gamma_i) \leq \gamma_i^{\max}$, where

$$\gamma_i^{\max} = \max_{\rho \in \mathbf{S}_\infty} \mathrm{Tr}(\Gamma_i \, \Pi\rho\Pi), \tag{38}$$

$$\gamma_i^{\min} = \min_{\rho \in \mathbf{S}_\infty} \mathrm{Tr}(\Gamma_i \, \Pi\rho\Pi). \tag{39}$$

For concreteness, we now specialize our discussion to the case where $[\Pi, \Gamma_i] = 0$ and $\Gamma_i \geq 0$ $\forall i$. This condition is satisfied for all the protocols we study in this paper. We emphasize that this is not a particularly strong assumption. With a judicious choice of Π, this condition can be achieved for many protocols. In the alternate, determining the bounds in Eqs. (38) and (39) without this assumption is also tractable. For example, tight bounds are derived even for noncommuting, nonpositive, and unbounded observables Γ_i in Ref. [25].

In the following theorem, we derive the desired bounds on expectations in the finite subspace.

Theorem 4. *Let* $\Gamma_i \geq 0$ *and* $[\Pi, \Gamma_i] = 0$. *If* $\mathrm{Tr}(\rho\bar{\Pi}) \leq W$ *and* $\mathrm{Tr}(\rho\Gamma_i) = \gamma_i$, *then* $\gamma_i - W\|\Gamma_i\|_\infty \leq \mathrm{Tr}(\Pi\rho\Pi\,\Gamma_i) \leq \gamma_i$.

Proof. By the commutation relation, it follows that

$$\mathrm{Tr}(\Pi\rho\Pi\Gamma_i) = \mathrm{Tr}\left(\sqrt{\Gamma_i}\Pi\rho\Pi\sqrt{\Gamma_i}\right) \tag{40}$$

$$= \mathrm{Tr}\left(\Pi\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\Pi\right). \tag{41}$$

We now find the upper and lower bounds separately.

For the upper bound, we simply note that the trace of a positive operator can only decrease under projection. Then,

$$\mathrm{Tr}\left(\Pi\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\Pi\right) \leq \mathrm{Tr}\left(\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\right) \tag{42}$$

$$= \gamma_i. \tag{43}$$

For the lower bound, recall Hölder's inequality, which states that for any two operators A and B, $\mathrm{Tr}(A^\dagger B) \leq \|A\|_p \|B\|_q$, where $1/p + 1/q = 1$ so that $\|\cdot\|_p$ and $\|\cdot\|_q$ are dual Schatten norms. Then,

$$\mathrm{Tr}\left(\Pi\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\Pi\right) = \mathrm{Tr}\left(\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\right)$$
$$- \mathrm{Tr}\left(\bar{\Pi}\sqrt{\Gamma_i}\rho\sqrt{\Gamma_i}\bar{\Pi}\right) \tag{44}$$

$$= \gamma_i - \mathrm{Tr}\left(\bar{\Pi}\rho\bar{\Pi}\Gamma_i\right) \tag{45}$$

$$\geq \gamma_i - \left\|\bar{\Pi}\rho\bar{\Pi}\right\|_1 \|\Gamma_i\|_\infty \tag{46}$$

$$\geq \gamma_i - W\|\Gamma_i\|_\infty, \tag{47}$$

where in Eq. (46) we use Hölder's inequality. Note that this lower bound is trivial for unbounded observables. ∎

Finally we consider the reduced state constraint $\mathrm{Tr}_B(\sigma) = \tau_A$. We could, using a complete Hermitian basis on system A, write the reduced state constraint as a set of expectations and expand it in the same manner as above. However, there are two better ways to perform the expansion. One approach works in general, while the other is tighter but only works in a specific case. In general, by the Fuchs–van de Graaf inequalities, $\frac{1}{2}\left\|\sigma - \tilde{\sigma}^\Pi\right\|_1 \leq \sqrt{2W - W^2}$. By the monotonicity of trace distance under channels, and by the fact that taking the partial trace is a channel, this implies the constraint $\frac{1}{2}\left\|\tau_A - \mathrm{Tr}_B(\tilde{\sigma}^\Pi)\right\|_1 \leq \sqrt{2W - W^2}$. Alternatively, in the case where the projection has the form $\Pi = \Pi_A \otimes \Pi_B$, a simple positivity argument implies the constraint $\mathrm{Tr}_B(\tilde{\sigma}^\Pi) \leq \Pi_A\tau_A\Pi_A$. This constraint can be used even when $\mathcal{H}_A$ is infinite dimensional.

Having expanded all the constraints in the case where the observables are positive operators and commute with

the projection, we summarize the definitions of the original infinite-dimensional optimization:

$$\underset{\rho}{\text{minimize}}: f(\rho)$$

$$\text{subject to}: \text{Tr}(\rho) = 1$$
$$\text{Tr}_B(\rho) = \tau_A \qquad (48)$$
$$\text{Tr}(\rho\Gamma_i) = \gamma_i$$
$$\rho \in \text{Pos}(\mathcal{H}_\infty)$$

and the expanded finite-dimensional optimization:

$$\underset{\tilde{\rho}}{\text{minimize}}: f(\tilde{\rho})$$

$$\text{subject to}: 1 - W \leq \text{Tr}(\tilde{\rho}) \leq 1$$
$$\frac{1}{2}\|\text{Tr}_B(\tilde{\rho}) - \tau_A\|_1 \leq \sqrt{2W - W^2} \qquad (49)$$
$$\gamma_i - W\|\Gamma_i\|_\infty \leq \text{Tr}(\tilde{\rho}\Gamma_i) \leq \gamma_i$$
$$\tilde{\rho} \in \text{Pos}(\mathcal{H}_N).$$

When $\Pi = \Pi_A \otimes \Pi_B$, the constraint on the reduced density matrix is instead $\text{Tr}_B(\tilde{\rho}^\Pi) \leq \Pi_A \tau_A \Pi_A$. By construction, $\mathbf{S}_N$ is a convex subset of $\tilde{D}(\mathcal{H}_N)$ containing $\Pi \mathbf{S}_\infty \Pi$. The tightness of the feasible set is guaranteed by the fact that for $W = 0$, $\mathbf{S}_N = \mathbf{S}_\infty$. For nonzero W, our numerical results in Sec. VII provide strong evidence that our definition of $\mathbf{S}_N$ is close to the optimal choice $\Pi \mathbf{S}_\infty \Pi$.

V. NUMERICAL IMPLEMENTATION

A. Review of reliable numerical method

Having formulated the finite-dimensional optimization of interest, we now wish to solve it numerically. In order to obtain a reliable lower bound in spite of floating-point imprecision and the imperfection of convex solvers, we use the numerical framework developed in Ref. [8]. We briefly summarize the two steps of the method here.

The first step is to approximately solve the optimization as written, using a method such as the Frank-Wolfe algorithm. This returns an approximate minimum, ρ_{opt}. The gradient of the objective function at this point $\nabla f(\rho_{\text{opt}})$ is computed. In the second step an expanded, linearized SDP is constructed from the first-step optimization. To linearize, the objective function is replaced with $\text{Tr}[\rho \nabla f(\rho_{\text{opt}})]$. Intuitively, this represents lower bounding the original convex function f by a tangent hyperplane. To expand, the feasible set is enlarged by a small amount to account for numerical imprecision. Finally, the dual of this expanded, linearized SDP is derived. This dual SDP is solved numerically and any feasible point is a reliable lower bound on the original convex minimization. As discussed in Appendix C, we make use of one small improvement in how we expand the set in the second

step, to account for numerical imprecisions, compared to Ref. [8]. This change is important because it allows us to improve our numerical results at long distances (see Sec. VI E).

Note that the constraint operators Γ_i and the POVM elements P_k defining f are, in general, infinite dimensional. However, as $\text{Tr}(\Pi\rho\Pi \ X) = \text{Tr}[(\Pi\rho\Pi)(\Pi X \Pi)]$, we can equivalently set $\Gamma_i \to \Pi\Gamma_i\Pi$ and $P_k \to \Pi P_k\Pi$. In the following discussion, we tacitly assume this substitution is made in order to represent the optimizations numerically.

B. SDP formulation for numerics

To apply the numerical framework, we need the feasible set of the convex minimization to be that of a SDP. To show that $\mathbf{S}_N$ is such a set, we rewrite the trace distance constraint for the reduced density matrix. The trace norm can be expressed as a SDP [36], which allows us to rewrite the constraint using slack variables. This equivalent reformulation of the finite-dimensional optimization is given by

$$\underset{\tilde{\rho},R,S}{\text{minimize}}: f(\tilde{\rho})$$

$$\text{subject to}: 1 - W \leq \text{Tr}(\tilde{\rho}) \leq 1$$
$$\text{Tr}(R) + \text{Tr}(S) \leq 2\sqrt{2W - W^2}$$
$$\text{Tr}_B(\tilde{\rho}) - R \leq \tau_A \qquad (50)$$
$$- S - \text{Tr}_B(\tilde{\rho}) \leq -\tau_A$$
$$\gamma_i - W\|\Gamma_i\|_\infty \leq \text{Tr}(\tilde{\rho}\Gamma_i) \leq \gamma_i$$
$$\tilde{\rho} \in \text{Pos}(\mathcal{H}_N)$$
$$R, S \in \text{Pos}(\mathcal{H}_A).$$

Let ξ denote the adjoint of the partial trace map Tr_B restricted to operators on $\mathcal{H}_N$. Following the numerical framework, the corresponding dual linearized SDP is

$$\underset{\vec{y},y_s,Y_1,Y_2}{\text{maximize}}: -\vec{y} \cdot (\vec{\gamma} + \vec{\epsilon}_{\text{rep}}) - y_s\left(2\sqrt{2W - W^2} + \epsilon'_{\text{rep}}\right)$$

$$- \text{Tr}(\tau_A Y_1) + \text{Tr}(\tau_A Y_2)$$

$$\text{subject to}: \sum_{i=1}^{2m} y_i\Gamma_i + \xi(Y_1) - \xi(Y_2) \geq -\nabla f(\rho_{\text{opt}})$$

$$y_s\mathbb{1}_A \geq Y_1$$
$$y_s\mathbb{1}_A \geq Y_2 \qquad (51)$$
$$\vec{y} \in \mathbb{R}^{2m}_{\geq 0}$$
$$y_s \in \mathbb{R}_{\geq 0}$$
$$Y_1, Y_2 \in \text{Pos}(\mathcal{H}_A),$$

where $\vec{\gamma} = (\{\gamma_i\}_{i=1}^m, \{W\|\Gamma_i\| - \gamma_i\}_{i=1}^m)$, $\vec{\Gamma} = (\{\Gamma_i\}_{i=1}^m, \{-\Gamma_i\}_{i=1}^m)$, and $\vec{\epsilon}_{\mathrm{rep}}$ and ϵ'_{rep} are the expansion parameters, which account for finite numerical precision (see Appendix C for further discussion).

As noted in Sec. IV D, the finite-dimensional optimization has a slightly different form when the projection only acts on Bob's system, that is $\Pi = \mathbb{1}_A \otimes \Pi_B$. In this case, the dual SDP is the same as above, but the term in the objective function and both constraints involving y_s are removed, Y_2 is set to zero, and the term $-\epsilon'_{\mathrm{rep}}\mathrm{Tr}(Y_1)$ is added to the objective function.

VI. DISCRETE-MODULATED CVQKD

We apply our approach for infinite-dimensional key-rate calculations to DMCVQKD. We begin by reviewing the protocol and setting up the infinite-dimensional optimization. We then apply the steps of our method from Sec. IV to relate the key rate to a finite-dimensional optimization. Finally, we solve this finite problem numerically to obtain key rates. Unless otherwise stated, all discussion is in the ideal detector scenario. We note that although continuity bounds have been considered in the context of DMCVQKD in Ref. [37], this is very different from our work, as the bounds in Ref. [37] are used to quantify how well Gaussian modulation is approximated.

A. Protocol description

We briefly review the DMCVQKD protocol. More details can be found in Ref. [23]. In each round of the quantum phase, Alice prepares one of d coherent signal states $\{|\alpha_i\rangle\}_{i=0}^{d-1}$ with probability $p(i)$. Our security proof works for any constellation of signal states, but we focus on the protocol with four symmetrically modulated states, also known as *quadrature phase-shift keying*. In this case, Alice prepares the states $\{|\alpha\rangle, |i\alpha\rangle, |-\alpha\rangle, |-i\alpha\rangle\}$ uniformly at random, for a fixed signal-state amplitude α. Bob then performs either a homodyne or heterodyne measurement. We focus on the case where he does a heterodyne measurement. Bob's POVM is then $\{1/\pi |\zeta\rangle \langle\zeta|\}_{\zeta\in\mathbb{C}}$, while Alice's POVM is $\{|i\rangle\langle i|\}_{i=0}^{d-1}$ as we work in the source-replacement picture.

We focus on reverse reconciliation as it is known to outperform direct reconciliation in terms of transmission distance. Again, our security proof can be easily adapted to direct reconciliation. In reverse reconciliation, Bob maps his heterodyne measurement outcome ζ to a key symbol based on which region in the complex plane, or phase space, the outcome falls in. The key map is visualized in Fig. 2. Bob can also perform postselection (sifting) to improve the key rate of the protocol [16].

After a large number of rounds, Alice and Bob move on to the classical phase. For parameter estimation, Alice
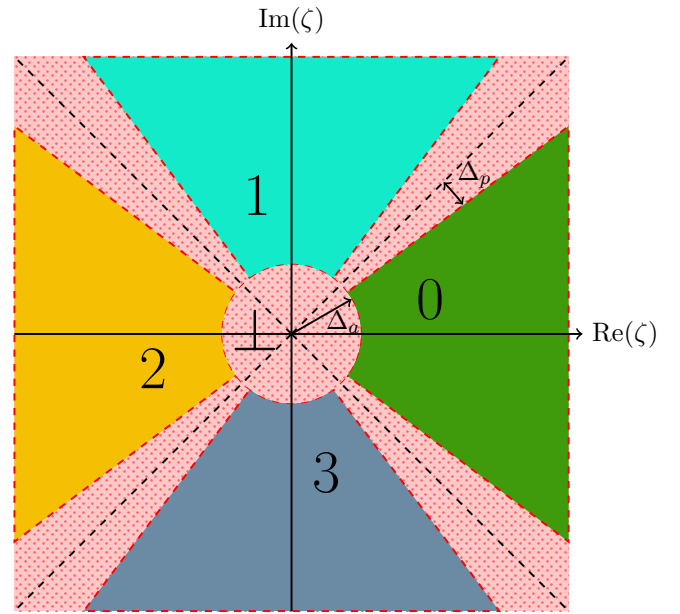

FIG. 2. Phase-space regions for the key map, with postselection, for reverse reconciliation in QPSK DMCVQKD. Bob obtains the measurement result ζ from the heterodyne detector. He maps the outcome to the symbol of the region containing ζ. Δ_a and Δ_p are amplitude and phase postselection parameters, respectively. The $\perp$ region corresponds to signals that are discarded.

announces which signal state was sent. Since Bob's measurement determines the Husimi Q-function, he could, in principle, perform complete tomography to uniquely determine the received conditional states. This would not extend reasonably to a finite-key analysis. We thus choose to only use certain coarse-grained observables as constraints. (It is then unnecessary for Bob to announce his fine-grained measurement results, he only needs to announce these coarse-grained expectations.) In Appendix E, we outline how to calculate the coarse-grained expectations from the fine-grained probability distribution. For reverse reconciliation with postselection, Bob also announces which signals are discarded. Finally, Alice and Bob perform error correction and privacy amplification on the sifted key.

B. Infinite-dimensional optimization

To formalize this description of the protocol, we define the objective function f and the observables Γ_i.

1. Objective function

Recall the definition of the postprocessing map in Eq. (4). Outside of the testing rounds, the only announcements in this protocol are for sifting. Since Bob defines the key map and makes the sifting decision, this means that Alice's data is irrelevant for the postprocessing map. Thus, for the purpose of the postprocessing map, we can take Alice's

POVM to trivially be $\{\mathbb{1}_A\}$. For Bob's POVM, we do not need to work with the fine-grained set and can instead deal directly with the coarse-grained region operators. These region operators can be expressed as linear combinations of the fine-grained POVM elements. The region operators corresponding to nondiscarded signals are

$$R_B^j = \frac{1}{\pi} \int_{\Delta_a}^{\infty} \int_{\frac{(2j-1)\pi}{4}+\Delta_p}^{\frac{(2j+1)\pi}{4}-\Delta_p} r \left| re^{i\theta} \right\rangle \left\langle re^{i\theta} \right| d\theta dr, \quad (52)$$

for $j \in \{0, 1, 2, 3\}$. Here Δ_a and Δ_p are postselection parameters for the amplitude and phase. Not performing postselection corresponds to setting $\Delta_a, \Delta_p = 0$. The key map g trivially maps Bob's private measurement result to the key register. Thus, we can omit the key map isometry altogether, and simply relabel the register $\bar{B}$ to Z. The simplified completely positive and trace the nonincreasing form of the postprocessing map Φ in Eq. (4) is then

$$\Phi(\rho_{ABE}) = \sum_{z=0}^{3} |z\rangle \langle z|_Z \otimes \text{Tr}_{AB}[\rho_{ABE}(\mathbb{1}_A \otimes R_B^z)], \quad (53)$$

where we omit trivial or redundant registers [23].

2. Observables

Alice and Bob's POVMs are as given above. As discussed, we need to choose a set of coarse-grained observables for Bob's side of parameter estimation. Typically, these observables are taken to be the quadratures and their higher moments [23]. However, we introduce a different set of observables, which will simplify our development, as they will commute with the projection we later define [Eq. (58)]. They will be parametrized by a list of complex numbers $\{\beta_i\}_{i=0}^{d-1}$. For any operator X, we use the shorthand notation $X_\gamma \equiv \hat{D}(\gamma) X \hat{D}^\dagger(\gamma)$, where $\hat{D}(\gamma)$ is the displacement operator with complex parameter γ. Denoting the photon-number operator by $\hat{n} \equiv \hat{a}^\dagger \hat{a}$, we choose, as constraint observables for our protocol, the set

$$\{\Gamma_i\} = \{|i\rangle \langle i| \otimes \hat{n}_{\beta_i}, |i\rangle \langle i| \otimes \hat{n}_{\beta_i}^2\}_{i=0}^{d-1}. \quad (54)$$

We consider second-order constraints in $\hat{n}$ as this will be necessary to make the weight W outside the subspace sufficiently small. We choose the displacements to be

$$\beta_i = \sqrt{\eta} \alpha_i, \quad (55)$$

which are the amplitudes of the coherent signal states after passing through a Gaussian channel with loss η. As discussed in Sec. VI E, our choice here is based on the expected channel behavior in an honest implementation of the protocol. We emphasize however that our security proof method works for any choice of $\{\beta_i\}$.

This choice of coarse-grained observables is of interest as it elucidates some of the essential working principles of the DMCVQKD protocol. The observables measure how spread out a state is compared to the coherent state $|\beta_i\rangle$. Intuitively, this characterizes the deviation from a generalized beam-splitting attack [38]. These constraints also dovetail with a natural choice for the finite subspace.

Note that although Bob's observables have a dependence on the signal state, Bob physically performs the same heterodyne measurement each round. For parameter estimation, he simply holds all fine-grained data and coarse grains only after Alice announces which signal state was sent.

3. Optimization formulation

We are now able to write down the infinite-dimensional optimization for DMCVQKD. Let $\mathcal{H}_A$ be the Hilbert space with dimension equal to the number of signal states d. Let $\mathcal{H}_B$ be the Hilbert space of a single optical mode, spanned by Fock states $\{|n\rangle\}_{n=0}^{\infty}$. The minimization is then

$$\underset{\rho}{\text{minimize}} : f(\rho)$$

$$\text{subject to} : \text{Tr}(\rho) = 1$$

$$\text{Tr}_B(\rho) = \sum_{i,j} \sqrt{p(i)p(j)} \langle \alpha_j | \alpha_i \rangle |i\rangle \langle j|$$

$$\text{Tr}\left[\rho\left(\tfrac{1}{p(i)}|i\rangle\langle i| \otimes \hat{n}_{\beta_i}\right)\right] = \langle \hat{n}_{\beta_i} \rangle \quad (56)$$

$$\text{Tr}\left[\rho\left(\tfrac{1}{p(i)}|i\rangle\langle i| \otimes \hat{n}_{\beta_i}^2\right)\right] = \langle \hat{n}_{\beta_i}^2 \rangle$$

$$\rho \in \text{Pos}(\mathcal{H}_A \otimes \mathcal{H}_B).$$

C. Finite-dimensional optimization and correction term

We apply the steps of our method to convert the preceding infinite-dimensional optimization into a tractable finite-dimensional one, and determine the associated correction term.

1. Choose subspace $\mathcal{H}_N$

Recall our general principle is to choose the subspace containing the most weight of the state under a typical channel model. As discussed in Sec. VI E, the channel model for a fiber-based implementation of this protocol is a lossy and noisy Gaussian channel. As noted above, for a pure-loss channel the expected coherent state is $|0_{\beta_i}\rangle$, where $|n_\gamma\rangle \equiv \hat{D}(\gamma)|n\rangle$. For the specific noise model we consider, the expected state will be a displaced thermal state. This means some weight leaks into displaced Fock states with $n > 0$. For the ith signal state, the best

projection on $\mathcal{H}_B$ is thus

$$\Pi^N_{B_{\beta_i}} = \sum_{n=0}^{N} |n_{\beta_i}\rangle\langle n_{\beta_i}|. \tag{57}$$

We refer to N as the subspace dimension parameter. The projection operator on the total Hilbert space is

$$\Pi^N \equiv \sum_{i=0}^{d-1} |i\rangle\langle i|_A \otimes \Pi^N_{B_{\beta_i}}. \tag{58}$$

Note that this projection commutes with the observables [Eq. (54)] because $\Pi^N_{B_{\beta_i}}$ commutes with $\hat{n}_{\beta_i}$. We use this fact to define the finite set (see Sec. VI C 4).

Unlike the truncated Fock basis considered in previous work [22,23], our finite subspace contains the full weight of the state when the channel is purely lossy. This is important as it ensures our numerics exactly reproduces the analytically solvable loss-only case.

2. Bound weight W outside subspace

To bound the weight outside the subspace, we analytically solve the dual of the SDP in Eq. (20). Our result is stated in the following theorem.

Theorem 5 (Bound on W for DMCVQKD). *For the DMCVQKD protocol, with Π^N as defined in Eq. (58), the weight outside the subspace is bounded by*

$$W = \sum_{i=0}^{d-1} p(i) \frac{\langle \hat{n}^2_{\beta_i}\rangle - \langle \hat{n}_{\beta_i}\rangle}{N(N+1)}. \tag{59}$$

Proof. To prove this theorem, we consider Bob's conditional states $\rho^i_B = [1/p(i)]\mathrm{Tr}_A[\rho_{AB}(|i\rangle\langle i|_A \otimes \mathbb{1}_B)]$. Let $\bar{\Pi}^N_{B_{\beta_i}} \equiv \mathbb{1}_B - \Pi^N_{B_{\beta_i}}$ and let $W_i \equiv \mathrm{Tr}(\rho^i_B \bar{\Pi}^N_{B_{\beta_i}})$ be the weight of the ith conditional state. We first show that $W = \sum_i p(i)W_i$.

$$\mathrm{Tr}(\rho\bar{\Pi}^N) = \mathrm{Tr}\left[\rho\left(\sum_{i=0}^{d-1} |i\rangle\langle i|_A \otimes \bar{\Pi}^N_{B_{\beta_i}}\right)\right] \tag{60}$$

$$= \sum_{i=0}^{d-1} \mathrm{Tr}\left(\tilde{\rho}^i_B \bar{\Pi}^N_{B_{\beta_i}}\right) \tag{61}$$

$$= \sum_{i=0}^{d-1} p(i)W_i. \tag{62}$$

Now, we only need to bound the weight of each conditional state. Using the constraints from Eq. (56), each of these

bounds can be expressed as a primal SDP.

$$\begin{aligned}
\underset{\rho}{\text{maximize}}&: \mathrm{Tr}\left(\bar{\Pi}^N_{B_{\beta_i}}\rho\right)\\
\text{subject to}&: \mathrm{Tr}(\rho) = 1\\
&\quad \mathrm{Tr}(\hat{n}_{\beta_i}\rho) = \langle\hat{n}_{\beta_i}\rangle\\
&\quad \mathrm{Tr}(\hat{n}^2_{\beta_i}\rho) = \langle\hat{n}^2_{\beta_i}\rangle\\
&\quad \rho \in \mathrm{Pos}\,(\mathcal{H}_B).
\end{aligned} \tag{63}$$

In order to find an upper bound on this primal SDP we consider its dual. By weak duality, it holds that a feasible solution to the dual SDP upper bounds the primal. In fact, strong duality holds for this SDP, so this upper bound can be made tight.

$$\begin{aligned}
\underset{\vec{y}}{\text{minimize}}&: y_1 + \langle\hat{n}_{\beta_i}\rangle y_2 + \langle\hat{n}^2_{\beta_i}\rangle y_3\\
\text{subject to}&: y_1\mathbb{1}_B + y_2\hat{n}_{\beta_i} + y_3\hat{n}^2_{\beta_i} - \bar{\Pi}^N_{B_{\beta_i}} \geq 0\\
&\quad \vec{y} \in \mathbb{R}^3.
\end{aligned} \tag{64}$$

A feasible solution for the dual is $y_1 = 0$, $y_2 = -1/N(N+1)$, $y_3 = 1/N(N+1)$. (This is in fact the optimal solution.) This can be easily verified as all operators in the constraint are diagonal in the $|n_{\beta_i}\rangle$ basis, so positivity is implied if and only if the diagonal entries are nonnegative. This solution leads to the objective value $W_i = \left(\langle\hat{n}^2_{\beta_i}\rangle - \langle\hat{n}_{\beta_i}\rangle\right)/N(N+1)$. Substituting into Eq. (62), the proof is complete. ∎

3. Determine correction term Δ

The correction term, as a function of the weight W and key map register dimension $|Z|$, is given in Theorem 2. We use the value of W determined in Sec. VI C 2. Regardless of whether postselection is performed, $|Z| = 4$ as there are four nondiscarded key outcomes (see Sec. IV C).

4. Choose finite set $\mathbf{S}_N$

By design, the projection Π^N [Eq. (58)] commutes with the positive observables Γ_i [Eq. (54)]. We can thus use the form of $\mathbf{S}_N$ in Eq. (49). The finite-dimensional

optimization is

$$\underset{\tilde{\rho}}{\text{minimize}} : f(\tilde{\rho})$$

$$\text{subject to: } 1 - W \le \text{Tr}(\tilde{\rho}) \le 1$$

$$\frac{1}{2} \|\text{Tr}_B(\tilde{\rho}) - \tau_A\|_1 \le \sqrt{2W - W^2}$$

$$\text{Tr}\left[\tilde{\rho}\left(\frac{1}{p(i)}|i\rangle\langle i| \otimes \hat{n}_{\beta_i}\right)\right] \le \langle \hat{n}_{\beta_i} \rangle \tag{65}$$

$$\text{Tr}\left[\tilde{\rho}\left(\frac{1}{p(i)}|i\rangle\langle i| \otimes \hat{n}_{\beta_i}^2\right)\right] \le \langle \hat{n}_{\beta_i}^2 \rangle$$

$$\tilde{\rho} \in \text{Pos}(\mathcal{H}_N),$$

where $\tau_A = \sum_{ij} \sqrt{p(i)p(j)} \langle \alpha_j | \alpha_i \rangle |i\rangle \langle j|$ and $W = \sum_i p(i)$ $\left(\langle \hat{n}_{\beta_i}^2 \rangle - \langle \hat{n}_{\beta_i} \rangle\right)/N(N+1)$. Note that the lower bounds in Eq. (49) are not useful as the observables Γ_i are unbounded.

In order to implement the optimization numerically, we need to choose a basis in which to represent our operators. The natural choice is $\{|i\rangle_A \otimes |n_{\beta_i}\rangle_B\}$. We compute the matrix elements of the observables and the objective function POVM in this basis in Appendix B. As this basis is not a standard construction for a bipartite Hilbert space, calculating the partial trace and its adjoint is slightly involved. We present these technical details in Appendix B.

D. DMCVQKD with trusted detector noise

In the previous section, we give a security proof for DMCVQKD assuming Bob's detector is ideal. We can extend this result to the scenario where Bob has imperfect but characterized detectors. We consider the specific model for detector imperfections given in Ref. [39]. Namely, the two homodyne detectors comprising the heterodyne measurement have an associated efficiency and electronic noise. To illustrate our approach, we focus on the case where the two detectors have the same efficiency η_d and electronic noise ν_{el}. In this case, Bob's POVM elements are displaced thermal states, as opposed to coherent states [39].

The protocol description is exactly the same as for the ideal detector case, except Bob's heterodyne detection is noisy. This noisy POVM enters into the optimization in two ways: changing the objective function f and the observables Γ_i. To evaluate the objective function defined by a noisy POVM f^{noisy}, we need only calculate the new matrix elements. This calculation is presented in Appendix D. Critically, due to our coarse graining, there is a simple relation between the noisy and ideal observables. Thus, our results on the bound of W carry over from the ideal detector scenario. For any operator A, we denote its noisy counterpart by $[A]'$. As shown in Appendix D, the ideal and noisy observables are related by

linear combinations,

$$\left[\hat{n}_{\sqrt{\eta_d}\beta_i}\right]' = \eta_d \hat{n}_{\beta_i} + \nu_{\text{el}}\mathbb{1}, \tag{66}$$

$$\left[\hat{n}_{\sqrt{\eta_d}\beta_i}^2\right]' = \eta_d^2 \hat{n}_{\beta_i}^2 + \eta_d(4\nu_{\text{el}} + 1 - \eta_d)\hat{n}_{\beta_i} + (2\nu_{\text{el}}^2 + \nu_{\text{el}})\mathbb{1}. \tag{67}$$

Bob measures the observables displaced by $\sqrt{\eta_d}\beta_i$. With these noisy expectations, the ideal ones can effectively be recreated by inverting the relationships in Eqs. (66), (67). Explicitly,

$$\langle \hat{n}_{\beta_i} \rangle^{\text{eff}} = \frac{\left\langle \left[\hat{n}_{\sqrt{\eta_d}\beta_i}\right]' \right\rangle - \nu_{\text{el}}}{\eta_d}, \tag{68}$$

$$\langle \hat{n}_{\beta_i}^2 \rangle^{\text{eff}} = \frac{1}{\eta_d^2}\Bigg[\left\langle \left[\hat{n}_{\sqrt{\eta_d}\beta_i}^2\right]' \right\rangle - 2\nu_{\text{el}}^2 - \nu_{\text{el}}$$

$$- (4\nu_{\text{el}} + 1 - \eta_d)\left(\left\langle \left[\hat{n}_{\sqrt{\eta_d}\beta_i}\right]' \right\rangle - \nu_{\text{el}}\right)\Bigg]. \tag{69}$$

The finite-dimensional optimization is then

$$\underset{\tilde{\rho}}{\text{minimize}} : f^{\text{noisy}}(\tilde{\rho})$$

$$\text{subject to: } 1 - W \le \text{Tr}(\tilde{\rho}) \le 1$$

$$\frac{1}{2} \|\text{Tr}_B(\tilde{\rho}) - \tau_A\|_1 \le \sqrt{2W - W^2}$$

$$\text{Tr}\left[\tilde{\rho}\left(\frac{1}{p(i)}|i\rangle\langle i| \otimes \hat{n}_{\beta_i}\right)\right] \le \langle \hat{n}_{\beta_i} \rangle^{\text{eff}} \tag{70}$$

$$\text{Tr}\left[\tilde{\rho}\left(\frac{1}{p(i)}|i\rangle\langle i| \otimes \hat{n}_{\beta_i}^2\right)\right] \le \langle \hat{n}_{\beta_i}^2 \rangle^{\text{eff}}$$

$$\tilde{\rho} \in \text{Pos}(\mathcal{H}_N),$$

where $\tau_A = \sum_{ij} \sqrt{p(i)p(j)} \langle \alpha_j | \alpha_i \rangle |i\rangle \langle j|$ and $W = \sum_i p(i)$ $\left[\left(\langle \hat{n}_{\beta_i}^2 \rangle^{\text{eff}} - \langle \hat{n}_{\beta_i} \rangle^{\text{eff}}\right)/N(N+1)\right]$.

E. Simulation results

1. Simulation parameters

To understand the performance of this protocol and demonstrate our security proof approach, we simulate expectation values obtained from a typical experiment. In particular, we model the signal states as passing through a noisy and lossy Gaussian channel. The transmittance η is modeled as a function of distance d according to $\eta = 10^{-k \cdot d/10}$, where k is the attenuation factor of the channel. We use a typical value for commercial-grade fiber $k = 0.2$ dB/km. The excess noise ξ is taken to be fixed at the channel input, for example as preparation noise, so that Bob sees the effective noise $\delta = \eta\xi$. The expectation values

for this simulation are $\langle \hat{n}_{\beta_i} \rangle = \delta/2$ and $\langle \hat{n}_{\beta_i}^2 \rangle = \delta(1 + \delta)/2$, as derived in Appendix E. This implies $W = \delta^2/[2N(N + 1)]$. We emphasize that our security proof does not depend on these parameter choices and simulation model, which are only used to illustrate the performance of the protocol in a typical implementation.

To account for realistic error-correction costs, $\delta_{\mathrm{EC}}^{\mathrm{leak}}$ in Eq. (8) is taken to be $H(Z) - \beta_{\mathrm{EC}} I(Z : X)$, where β_{EC} characterizes the error-correction efficiency. We use $\beta_{\mathrm{EC}} = 0.95$ as a representative value. The error-correction cost is calculated by simulating the joint probability distribution obtained by Alice and Bob (see Appendix E).

All our algorithms are implemented in MATLAB R2019B, using the convex optimization package CVX 2.1 [40,41] with the MOSEK 8.0.0.60 solver [42]. The Frank-Wolfe algorithm, with a maximum of 30 iterations, is used to solve the first step of the numerical method. All parameter optimizations use the fminbnd algorithm included in the MATLAB distribution, which uses a combination of parabolic interpolation and golden-section search.

We emphasize that all key rate plots include the correction term unless stated otherwise. That is, $R^{\infty} = C_{\mathrm{num}} - \delta_{\mathrm{EC}}^{\mathrm{leak}} - \Delta(W)$, where C_{num} denotes the reliable numerical lower bound on $f(\tilde{\rho}^N)$. In order to evaluate the effect of the correction term, and to compare with previous work using the photon-number cutoff assumption, we find it useful to consider the uncorrected values, defined as $C_{\mathrm{num}} - \delta_{\mathrm{EC}}^{\mathrm{leak}}$.

2. Key-rate plots

We present key-rate plots for different choices of channel parameters η and ξ, protocol parameters α, Δ_a, and Δ_p, and the subspace dimension parameter N. For the trusted noise scenario, we also consider η_d and ν_{el}.

In Fig. 3, we compare the key rates and uncorrected values from our dimension reduction method to the key rates under the photon-number cutoff assumption obtained in Ref. [23]. In order to enable a meaningful comparison, we use the protocol parameters from Ref. [23]. The uncorrected values, which are equal to the key rate before subtracting the correction term Δ, are essentially identical to those in Ref. [23]. As the results from Ref. [23] are an upper bound on the key rate, this indicates our choice of $\mathbf{S}_N$ is tight. Further, our corrected key rates are very close to the uncorrected values. This illustrates our correction term is small for reasonable values of the subspace dimension parameter N, at low channel excess noise (see Fig. 5).

As our key rates are very similar to the ones under the cutoff assumption, the qualitative conclusions of previous work [23,39] are confirmed to hold under our precise treatment, without the previous working assumptions. We thus defer to Ref. [23,39] for more extensive parameter exploration, and only focus on some important results in this section.
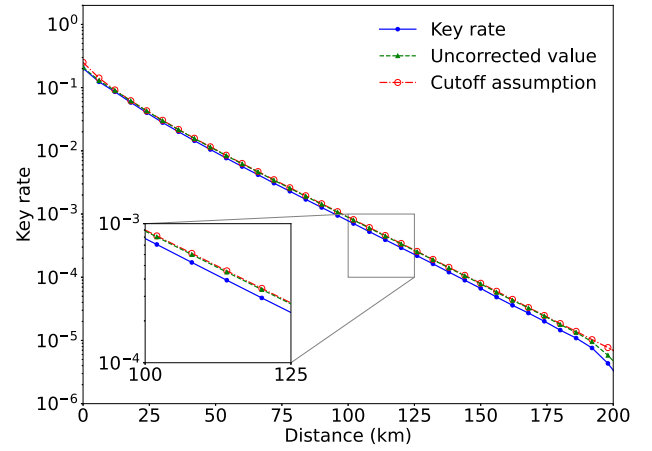

FIG. 3. Comparison of key rates and uncorrected values from our dimension reduction method with key rates under the photon-number cutoff assumption from Ref. [23]. Results are plotted versus distance with excess noise $\xi = 0.01$, and are in the ideal detector scenario. Postselection parameters and signal-state intensities from Ref. [23] are used: $\alpha = 0.6$, $\Delta_p = 0$, and Δ_a is optimized with a coarse-grained search over $[0.5, 0.65]$. The subspace dimension parameter is $N = 20$.

In Fig. 4, we plot the ideal detector key rates for different channel parameters. The signal-state intensity and post-selection parameters are numerically optimized for each distance and value of excess noise, using $N = 10$. The key rates are calculated using $N = 40$, except for a small number of points where we use $N = 30$ to ameliorate numerics issues, as discussed in Appendix C. Using postselection extends the range of the protocol for high excess noise while also reducing the amount of data processing for error correction, which can be a bottleneck in actual implementations. For example, for 2% excess noise, postselection increases the maximum distance by around 50 km, while discarding 40% of the signals. The small number of outlying points that deviate from the trend are due to numerical issues inherent to convex solvers. We emphasize that these key rates are still rigorous, and can be improved by using higher numerical precision.

To illustrate the relative size of the correction term, we plot it as a fraction of the uncorrected value in Fig. 5, at a fixed distance of 15 km. More precisely, for each value of excess noise ξ and subspace dimension parameter N, we plot the fractional correction term $\Delta(W)/C_{\mathrm{num}}$, where $W = (\eta\xi)^2/[2N(N + 1)]$, while C_{num} is computed at a fixed value of $N = 40$ for each ξ. (In the range considered here, C_{num} has a negligible dependence on N.) The protocol parameters are the same as in Fig. 4. We see that for a pure-loss channel, the correction term is zero since our subspace fully contains the simulation state, as discussed in Sec. VI C 1. For small values of excess noise, the correction term is negligible even for small N. For larger values of excess noise, and especially in the high loss
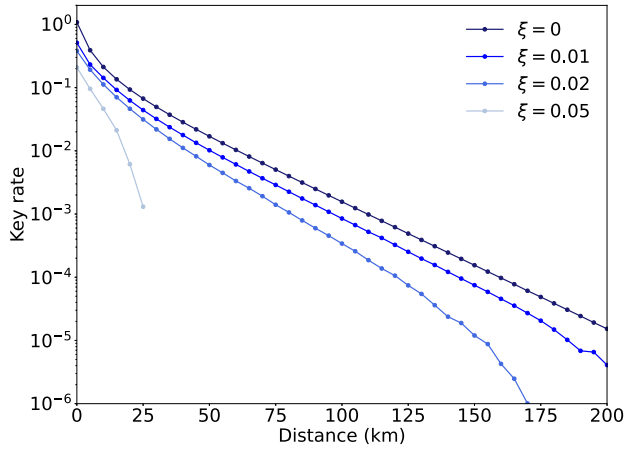
FIG. 4. Ideal detector secure key rates versus transmission distance, for different values of excess noise ξ, with optimized postselection parameters and signal-state intensity. For each point, the better of the two results from $N = 30$ and $N = 40$ is used, with the majority of points from $N = 40$. Postselection can improve the noise tolerance and range of the protocol.

FIG. 6. Optimal signal-state amplitude α_{opt} versus transmission distance, for different values of excess noise ξ and in the ideal detector scenario. The amplitude is optimized in the range $[0.5, 2]$, with $\Delta_a = \Delta_p = 0$ and $N = 10$.

regime, N must be increased to obtain reasonable results. This is because the correction term scales like the loss η for all values of excess noise ξ, while, for nonzero ξ, the key rate scales worse than η. It is an interesting avenue for future research to determine if a smaller correction term can be found.

The optimal signal-state amplitudes α_{opt} are shown in Fig. 6. The optimization range $[0.5, 2]$ is sufficient for almost all parameter choices, though α_{opt} tends to infinity as distance and excess noise tend to zero. The general trend is that α_{opt} decreases as distance and excess noise increase. With α_{opt} fixed, Δ_a is optimized over $[0, 1]$. (One could jointly optimize all protocol parameters, but we do
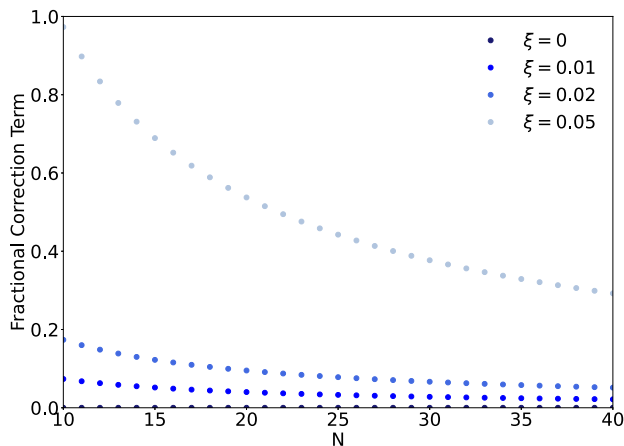

FIG. 5. The fractional correction term versus the subspace dimension parameter N for different values of excess noise, in the ideal detector scenario. The distance is 15 km and protocol parameters are optimized.
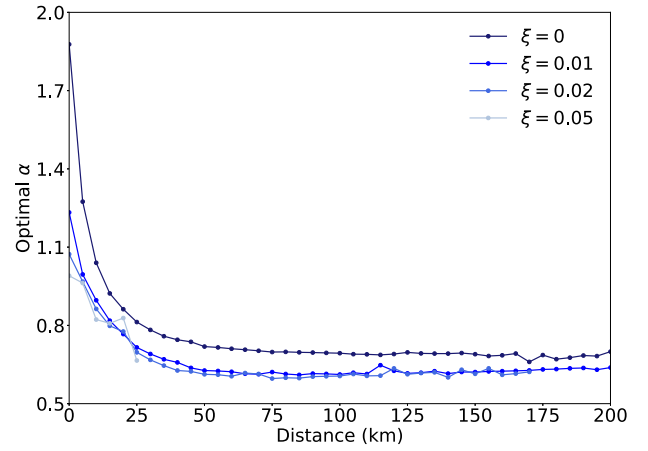
not expect this to noticeably improve the key rates.) We find that the optimal value for Δ_p seems to always be zero, so phase postselection is omitted altogether. Thus, while the postselection pattern used here is a simple and intuitive one, it is an interesting future research topic to investigate other postselection patterns.

In Fig. 7, we consider the key rates in the trusted detector noise scenario. We take the channel to have 1% excess noise, and consider different values of detector efficiency and electronic noise. The protocol parameters are optimized for the ideal detector scenario, and the same parameters are used for each of the different trusted noise cases. We observe that even with large detector imperfections and 1% channel excess noise, it is possible to generate secure key at 200 km. As expected, trusted detector noise does not significantly alter the scaling of the key rates. This is markedly different from the effect of channel excess noise (see Fig. 4).

DMCVQKD is intended to be a more experimentally feasible alternative to Gaussian-modulated (GM) CVQKD [19,43]. It is thus of interest to compare the performance of the two protocols. We perform a very basic comparison in Fig. 8, using the same error-correction efficiency $\beta_{\mathrm{EC}} = 0.95$, detector loss $\eta_d = 0.6$, and electronic noise $\nu_{\mathrm{el}} = 0.05$. The signal variance is optimized for GM, while the optimized protocol parameters for DM are the same as in Fig. 7. We note that a complete and in-depth comparison of the two protocols would have to account for many more implementation details. For a pure-loss channel, the GM key rates are around an order of magnitude higher. At higher excess noise, the gap is larger as GM is more robust to channel noise. However, both protocols are largely unaffected by trusted detector imperfections. At $\xi = 0.01$, a typical value for channel excess noise, the key rates scale
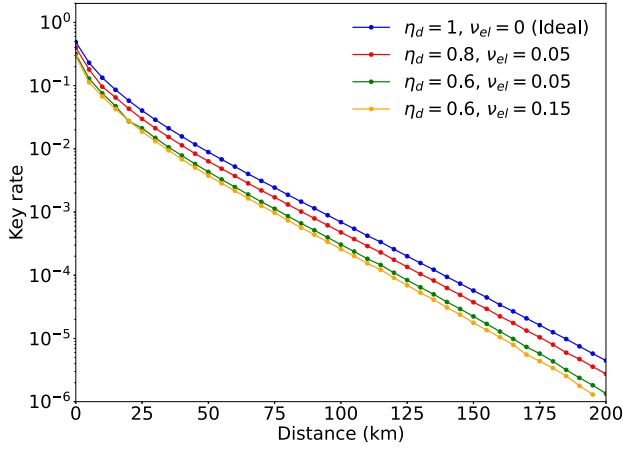
FIG. 7. Key rates versus distance for different trusted detector imperfections, with excess noise $\xi = 0.01$. Protocols are evaluated with the same optimized protocol parameters, including postselection, as in Fig. 4. The subspace dimension parameter is $N = 10$.

similarly. Further, we expect the tolerance of DM to channel noise can be improved by using a larger constellation of signal states.

VII. COMPARISON TO FLAG-STATE SQUASHER

Our dimension reduction (DR) approach encompasses another method known as the flag-state squasher (FSS) [12]. The FSS also obtains lower bounds on the key rate by solving a finite-dimensional optimization. However, the FSS is restricted to protocols where both the key map POVM elements and constraint observables commute with the projection. Notably, this is not the case for DMCVQKD. In this section, we compare our method to
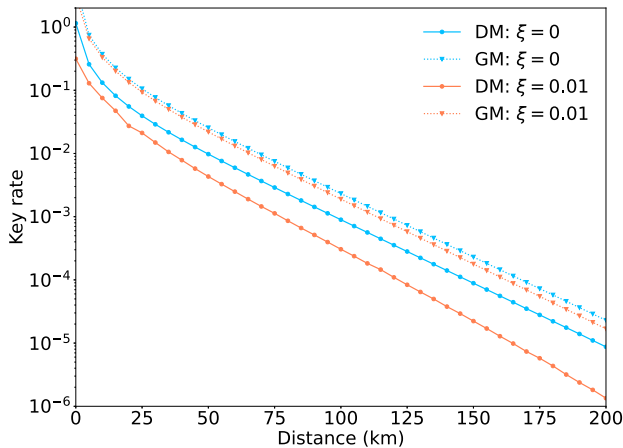

FIG. 8. Comparison of Gaussian [43] and discrete-modulation key rates for different values of excess noise ξ, in the trusted noise scenario with $\eta_d = 0.6$, $\nu_{el} = 0.05$. Parameters for both protocols are optimized. The subspace dimension parameter is $N = 10$.

the FSS both analytically and numerically. This demonstrates the advantages of our method and offers further insight into the FSS approach.

A. Analytical comparison

We briefly summarize the FSS, deferring a complete description to Ref. [12]. As usual, Alice's POVM is given by $\{|j\rangle \langle j|_A\}$, while Bob's POVM is $\{\Gamma_B^i\}$. The corresponding probabilities are $\{\gamma_{ij}\}$. The FSS also requires choosing a projection $\Pi = \mathbb{1}_A \otimes \Pi_B$ onto a finite subspace and upper bounding the weight W outside that subspace. It is assumed that $[\Gamma_B^i, \Pi_B] = 0$.

Define a flag Hilbert space $\mathcal{H}_F$, with dimension equal to the number of elements in Bob's POVM. The finite-dimensional optimization is over density matrices in $\mathcal{H}_N \oplus \mathcal{H}_F$. Bob's new POVM is $\tilde{\Gamma}^i = \Pi\Gamma^i\Pi \oplus |i\rangle \langle i|_F$. Alice's POVM and the expectation values are unchanged. The objective function is $f_{\text{FSS}}(\tilde{\rho}_N \oplus \tilde{\sigma}_F) = f(\tilde{\rho}_N)$, i.e., it simply discards the flag portion and evaluates the usual key-rate function on the remaining portion. This completes the formulation of the finite-dimensional optimization for the FSS.

The squashing map Λ_B is a channel from $\mathcal{H}_\infty$ to $\mathcal{H}_N \oplus \mathcal{H}_F$ defined as $\Lambda_B(\rho_\infty) = \Pi\rho_\infty\Pi \oplus \sum_i \text{Tr}_B[\rho_\infty(\Gamma^i - \Pi\Gamma^i\Pi)] |i\rangle \langle i|_F$. Note that ρ_∞ is feasible for the original infinite-dimensional optimization if and only if $\Lambda(\rho_\infty)$ is feasible for the finite-dimensional one. In this sense, we can think of the flag-state squasher as implicitly solving over the tightest possible choice of $\mathbf{S}_N$, namely $\Pi\mathbf{S}_\infty\Pi$.

As a special case, our method can be applied to any protocol admitting a flag-state squasher. The FSS requires choosing a projection Π and getting a bound on weight W, which establishes the first two steps of our method. Since the key map POVM elements commute with the projection, we can set $\Delta = 0$ (Theorem 3). All observables are POVM elements, so we can use the explicit form of $\mathbf{S}_N$ in Eq. (49). This establishes the last two steps of our method. We can now compare both approaches in the following theorem.

Theorem 6. *For a fixed projection Π and weight W, our dimension reduction (DR) method gives the same key rate as the flag-state squasher (FSS) when $\mathbf{S}_N = \Pi\mathbf{S}_\infty\Pi$.*

Proof. Let $\Lambda(\rho_\infty) = \tilde{\rho}_N \oplus \tilde{\sigma}_F$ be a state reaching the minimum in the FSS optimization. By definition, $f_{\text{FSS}}(\tilde{\rho}_N \oplus \tilde{\sigma}_F) = f_{\text{DR}}(\tilde{\rho}_N)$. By the definition of Λ, $\tilde{\rho}_N = \Pi\rho_\infty\Pi$. Thus, $\tilde{\rho}_N \in \mathbf{S}_N$, so is feasible for the dimension reduction optimization. Since this optimization is a minimization, $R_{\text{FSS}}^\infty \geq R_{\text{DR}}^\infty$.

Conversely, let $\tilde{\rho}_N$ be a subnormalized state reaching the minimum in the dimension reduction optimization. By the definition of $\mathbf{S}_N$, $\tilde{\rho}_N = \Pi\rho_\infty\Pi$ for some state $\rho_\infty \in \mathbf{S}_\infty$. By definition, $f_{\text{DR}}(\tilde{\rho}_N) = f_{\text{FSS}}[\Lambda(\rho_\infty)]$. By the property of

the squashing map, $\Lambda(\rho_\infty)$ is feasible for the FSS optimization. Since the FSS optimization is a minimization, $R_{DR}^\infty \geq R_{FSS}^\infty$. ∎

If $\mathbf{S}_N$ is not chosen optimally, then our dimension reduction method gives a lower key rate. In practice, our explicit prescription for choosing $\mathbf{S}_N$ in Eq. (49) gives very similar key rates to the flag state squasher (see Sec. VII B), suggesting this choice is essentially optimal.

In addition to being more general, our method has an important advantage compared to the flag-state squasher. Our finite-dimensional optimization is over a smaller Hilbert space, since we do not require flag-state dimensions. Therefore, if we compare fixed *total* dimension, which roughly determines the runtime, our method can give higher key rates than the flag-state squasher. Some protocols can have a very large number of POVM constraints. For the flag-state squasher, using all the constraints would make the runtime prohibitive, as the dimension of the problem depends on the number of constraints. Thus, a smaller set of coarse-grained POVM elements is typically used. Our dimension reduction method is not limited by the number of constraints and can thus handle the fine-grained POVM directly, potentially giving better key rates.

B. Numerical comparison: Unbalanced phase-encoded BB84

Having provided an analytical comparison between our method and the flag-state squasher, we now perform a sample numerical comparison of the two methods. We consider the unbalanced phase-encoded BB84 protocol, to which the flag-state squasher has recently been applied [13]. We defer a complete description of the protocol to Ref. [13]. Briefly, this is a phase-encoded BB84 protocol where Alice and Bob's interferometers each have a loss $1 - \kappa$ only in the arm with the phase modulator; hence the term unbalanced. The projection is $\Pi_B = \sum_{\substack{0 \leq n_1, n_2 \\ n_1 + n_2 \leq N}} |n_1, n_2\rangle \langle n_1, n_2|$, where $|n_1, n_2\rangle$ are two-mode Fock basis states. The weight W is bounded using the fact that the frequency of cross clicks increases with photon number [13].

In Fig. 9 we compare the key rates from our method and the flag-state squasher, for a channel with transmittance η and for different interferometer asymmetric transmittance κ. All parameters are the same as in Fig. 3(a) of Ref. [13], and the signal-state intensity is optimized separately for each method and parameter choice. We see that our method gives essentially identical key rates. In conjunction with Theorem 6, this provides strong numerical evidence that our heuristic choice of $\mathbf{S}_N$ in Eq. (49) is tight. While a more thorough benchmarking would be in order, we remark that for generating the data in Fig. 9, our method is approximately 5 times faster than the flag-state squasher
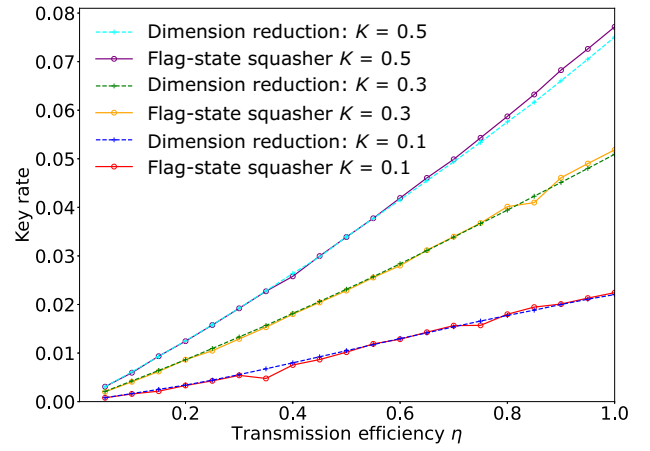

FIG. 9. Key rates for unbalanced phase-encoded BB84, versus transmission efficiency η, for different values of asymmetric interferometer loss $1 - \kappa$. It is clear that the key rates from our dimension reduction method are nearly identical to those from the flag-state squasher, indicating the tightness of our method in practice. In generating the data for this graph, our dimension reduction method is approximately 5 times faster than the flag-state squasher as implemented in Ref. [13].

as implemented in Ref. [13] (using the same SDPT3 solver [44,45]).

VIII. CONCLUSION

In summary, we establish a framework to lower bound a large dimensional convex optimization using a judiciously chosen smaller dimensional one. We show how this framework can be used to reduce the dimension of QKD key rate calculations. This allows existing numerical tools for finite-dimensional key rate calculations to be applied to protocols in infinite-dimensional Hilbert spaces. This allows us to do more detailed modeling of imperfections in devices. An important application of our method is to prove the asymptotic security of DMCVQKD with an arbitrary number of modulated states. As a concrete example, we apply this method to the quadrature phase-shift keying scheme in both ideal and trusted detector noise scenarios. We show that discrete modulation key rates can scale similarly to Gaussian modulation. Moreover, we rigorously demonstrate that postselection of data can improve the key rates for DMCVQKD. Using unbalanced phase-encoded BB84 as an example, we show that our approach can achieve key rates nearly identical to those from the flag-state squasher, while having an improved runtime.

Some directions for future work are as follows. One may be able to make the correction term for the QKD objective function smaller by using the commutation relations of the POVM elements and the projection. Qualitatively, this would interpolate between the two cases we consider. There may also be tighter ways to construct the finite set. This could involve using additional properties of the

constraint operators in specific cases. For DMCVQKD in particular, one may consider the effects of using more modulated states and different postselection regions in order to increase the key rate. Given the recent development of a finite-key numerical framework [46], we hope to extend our dimension reduction method to finite-key analysis of protocols in infinite-dimensional Hilbert spaces. We expect that key elements of the method, including bounding the weight outside the subspace and expanding the feasible set, will lift to the finite-key analysis.

APPENDIX A: UNIFORM CONTINUITY BOUND

Here we prove an extension of Lemma 2 in Ref. [33] to subnormalized states. Our development closely parallels that result. Although we are only interested in showing the conditional entropy is uniformly close to decreasing under projection with correction Δ, we effectively have to derive uniform continuity to determine Δ; so for completeness we give the overall uniform continuity bound as well. Note that the correction term in Eq. (A3) is smaller than Eq. (A2).

Theorem 7 (Uniform continuity and UCDUP of conditional entropy). *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two Hilbert spaces where the dimension of $\mathcal{H}_A$ is $|A|$ while $\mathcal{H}_B$ can be infinite dimensional. Let $\tilde{\rho}_{AB}, \tilde{\sigma}_{AB} \in \tilde{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be two subnormalized states; we omit the system subscripts for readability. Without loss of generality, suppose $Tr(\tilde{\rho}) \geq Tr(\tilde{\sigma})$. Let $\frac{1}{2}\|\tilde{\rho} - \tilde{\sigma}\|_1 \leq \epsilon \leq 1$, $\frac{1}{2}Tr(\tilde{\rho} - \tilde{\sigma}) = \delta$ and $\frac{1}{2}Tr(\tilde{\rho} + \tilde{\sigma}) = a$. Let $\epsilon' = \epsilon + \delta$ and $\epsilon'' = \epsilon - \delta$. Then, it holds that*

$$\left| H(A|B)_{\tilde{\rho}} - H(A|B)_{\tilde{\sigma}} \right| \leq 2\epsilon \log_2 |A|$$
$$+ (a + \epsilon) \max \left\{ h\left(\frac{\epsilon'}{a + \epsilon}\right), h\left(\frac{\epsilon''}{a + \epsilon}\right) \right\}. \quad (A1)$$

If $\tilde{\rho}$ and $\tilde{\sigma}$ are classical-quantum states, that is $\tilde{\rho} = \sum_{i=1}^{|A|} |i\rangle \langle i|_A \otimes \tilde{\rho}_B^i$ and $\tilde{\sigma} = \sum_{i=1}^{|A|} |i\rangle \langle i|_A \otimes \tilde{\sigma}_B^i$, then

$$\left| H(A|B)_{\tilde{\rho}} - H(A|B)_{\tilde{\sigma}} \right| \leq \epsilon' \log_2 |A|$$
$$+ (a + \epsilon) \max \left\{ h\left(\frac{\epsilon'}{a + \epsilon}\right), h\left(\frac{\epsilon''}{a + \epsilon}\right) \right\}, \quad (A2)$$

and

$$H(A|B)_{\tilde{\sigma}} - H(A|B)_{\tilde{\rho}} \leq \epsilon'' \log_2 |A| + (a + \epsilon) h\left(\frac{\epsilon''}{a + \epsilon}\right). \quad (A3)$$

Proof. We can assume $\frac{1}{2}\|\tilde{\rho} - \tilde{\sigma}\|_1 = \epsilon$ since our bound will be increasing in ϵ. Note that $\delta \leq \epsilon$. As usual, ρ and σ denote the normalized $\tilde{\rho}$ and $\tilde{\sigma}$. Let $\cdot_+$ denote the positive part of a Hermitian operator. The proof consists of a series of operator inequalities and applications of strong subadditivity.

We first determine the trace of the positive and negative parts of $\tilde{\rho} - \tilde{\sigma}$. To do this, consider the eigenvalues λ_i of $\tilde{\rho} - \tilde{\sigma}$. By assumption, $\sum |\lambda_i| = \|\tilde{\rho} - \tilde{\sigma}\|_1 = 2\epsilon$ and $\sum \lambda_i = Tr(\tilde{\rho} - \tilde{\sigma}) = 2\delta$. Thus, $Tr[(\tilde{\rho} - \tilde{\sigma})_+] = \sum_{\lambda_i \geq 0} \lambda_i = \epsilon + \delta = \epsilon'$. Similarly, $Tr[(\tilde{\rho} - \tilde{\sigma})_-] = -\sum_{\lambda_i < 0} \lambda_i = \epsilon - \delta = \epsilon''$.

Thus, $(1/\epsilon')(\tilde{\rho} - \tilde{\sigma})_+$ and $(1/\epsilon'')(\tilde{\rho} - \tilde{\sigma})_-$ are normalized states. Denote them by Δ and Δ', respectively. After some rearrangement, we can define a third state ω satisfying

$$\omega = \frac{Tr\tilde{\sigma}}{Tr\tilde{\sigma} + \epsilon'}\sigma + \frac{\epsilon'}{Tr\tilde{\sigma} + \epsilon'}\Delta$$
$$= \frac{Tr\tilde{\rho}}{Tr\tilde{\rho} + \epsilon''}\rho + \frac{\epsilon''}{Tr\tilde{\rho} + \epsilon''}\Delta'. \quad (A4)$$

Note that $Tr\tilde{\sigma} + \epsilon' = Tr\tilde{\rho} + \epsilon'' = a + \epsilon$. We find an upper and lower bound on $H(A|B)_\omega$, and combine them to get our final result.

The lower bound simply follows from the concavity of conditional entropy and the definition of ω in Eq. (A4),

$$H(A|B)_\omega \geq \frac{Tr\tilde{\sigma}}{a + \epsilon}H(A|B)_\sigma + \frac{\epsilon'}{a + \epsilon}H(A|B)_\Delta. \quad (A5)$$

For the upper bound, we first rewrite the conditional entropy in terms of the relative entropy as follows [47]:

$$-H(A|B)_{\omega_{AB}} = \min_{\xi_B} D(\omega_{AB} || \mathbb{1}_A \otimes \xi_B). \quad (A6)$$

Note that the minimum is achieved at $\xi_B = \omega_B = Tr_A(\omega_{AB})$. Expanding the definition of the relative entropy,

we have

$$H(A|B)_\omega = -D(\omega_{AB}||\mathbb{1}_A \otimes \omega_B) \quad (A7)$$

$$= H(\omega) + \mathrm{Tr}[\omega(\mathbb{1}_A \otimes \log_2 \omega_B)]. \quad (A8)$$

We upper bound the first term using strong subadditivity,

$$H(\omega) = H\left(\frac{\mathrm{Tr}\tilde\rho}{a+\epsilon}\rho + \frac{\epsilon''}{a+\epsilon}\Delta'\right) \quad (A9)$$

$$\leq \frac{\mathrm{Tr}\tilde\rho}{a+\epsilon}H(\rho) + \frac{\epsilon''}{a+\epsilon}H(\Delta') + h\left(\frac{\epsilon''}{a+\epsilon}\right). \quad (A10)$$

In the second term, we simply insert the definition of ω and expand. Thus, we have

$$H(A|B)_\omega \leq \frac{\mathrm{Tr}\tilde\rho}{a+\epsilon}H(\rho) + \frac{\epsilon''}{a+\epsilon}H(\Delta')$$
$$+ \frac{\mathrm{Tr}\tilde\rho}{a+\epsilon}\mathrm{Tr}[\rho(\mathbb{1}_A \otimes \log_2 \omega_B)]$$
$$+ \frac{\epsilon''}{a+\epsilon}\mathrm{Tr}[\Delta'(\mathbb{1}_A \otimes \log_2 \omega_B)] + h\left(\frac{\epsilon''}{a+\epsilon}\right) \quad (A11)$$

$$= -\frac{\mathrm{Tr}\tilde\rho}{a+\epsilon}D(\rho||\mathbb{1}_A \otimes \omega_B) - \frac{\epsilon''}{a+\epsilon}D(\Delta'||\mathbb{1}_A \otimes \omega_B)$$
$$+ h\left(\frac{\epsilon''}{a+\epsilon}\right), \quad (A12)$$

where we recombine the terms into relative entropies. We now use the relation in Eq. (A6) again, to obtain

$$H(A|B)_\omega \leq \frac{\mathrm{Tr}\tilde\rho}{a+\epsilon}H(A|B)_\rho + \frac{\epsilon''}{a+\epsilon}H(A|B)_{\Delta'}$$
$$+ h\left(\frac{\epsilon''}{a+\epsilon}\right). \quad (A13)$$

The upper and lower bounds on $H(A|B)_\omega$, in Eqs. (A13) and (A5), respectively, can be combined to obtain

$$\frac{\mathrm{Tr}\tilde\sigma}{a+\epsilon}H(A|B)_\sigma + \frac{\epsilon'}{a+\epsilon}H(A|B)_\Delta \leq \frac{\mathrm{Tr}\tilde\rho}{a+\epsilon}H(A|B)_\rho$$
$$+ \frac{\epsilon''}{a+\epsilon}H(A|B)_{\Delta'} + h\left(\frac{\epsilon''}{a+\epsilon}\right), \quad (A14)$$

$$H(A|B)_{\tilde\sigma} - H(A|B)_{\tilde\rho} \leq \epsilon'' H(A|B)_{\Delta'} - \epsilon' H(A|B)_\Delta$$
$$+ (a+\epsilon)h\left(\frac{\epsilon''}{a+\epsilon}\right). \quad (A15)$$

By repeating the proof but interchanging the two expressions for ω, we similarly obtain

$$H(A|B)_{\tilde\rho} - H(A|B)_{\tilde\sigma} \leq \epsilon' H(A|B)_\Delta - \epsilon'' H(A|B)_{\Delta'}$$
$$+ (a+\epsilon)h\left(\frac{\epsilon'}{a+\epsilon}\right). \quad (A16)$$

Conditional entropies of normalized states are bounded between $\pm \log_2 |A|$. Thus, we have

$$\left|H(A|B)_{\tilde\rho} - H(A|B)_{\tilde\sigma}\right| \leq 2\epsilon \log_2 |A|$$
$$+ (a+\epsilon)\max\left\{h\left(\frac{\epsilon'}{a+\epsilon}\right), h\left(\frac{\epsilon''}{a+\epsilon}\right)\right\}. \quad (A17)$$

When $\tilde\rho$ and $\tilde\sigma$ are both classical-quantum states, Δ and Δ' are also both classical-quantum states. Then, their conditional entropy is between 0 and $\log_2 |A|$. This gives the tighter bound of

$$\left|H(A|B)_{\tilde\rho} - H(A|B)_{\tilde\sigma}\right| \leq \epsilon' \log_2 |A| + (a+\epsilon)\max$$
$$\left\{h\left(\frac{\epsilon'}{a+\epsilon}\right), h\left(\frac{\epsilon''}{a+\epsilon}\right)\right\}. \quad (A18)$$

Similarly,

$$H(A|B)_{\tilde\sigma} - H(A|B)_{\tilde\rho} \leq \epsilon'' \log_2 |A| + (a+\epsilon)h\left(\frac{\epsilon''}{a+\epsilon}\right). \quad (A19)$$

∎

Corollary 1. *Let $\tilde\rho_{AB}$ and $\tilde\sigma_{AB}$ be two bipartite subnormalized classical-quantum states with $Tr(\tilde\rho) \geq Tr(\tilde\sigma)$; the dimension of system B can be infinite. Let $\frac{1}{2}\|\tilde\rho - \tilde\sigma\|_1 \leq \epsilon \leq 1$. Then,*

$$H(A|B)_{\tilde\sigma} - H(A|B)_{\tilde\rho} \leq \epsilon \log_2 |A| + (1+\epsilon)h\left(\frac{\epsilon}{1+\epsilon}\right). \quad (A20)$$

Proof. Begin with the third statement of Theorem 7. We can upper bound ϵ'' in the first term on the right-hand side by ϵ. Then, since the function $g(a) = (a+\epsilon)h[c/(a+\epsilon)]$ is increasing on $a \in [0, 1]$, we can upper bound the second term on the right-hand side by evaluating it at $a = 1$. We have

$$\frac{\epsilon''}{1+\epsilon} \leq \frac{\epsilon}{1+\epsilon} \leq \frac{1}{2}. \quad (A21)$$

Since the binary entropy is increasing on $[0, \frac{1}{2}]$,

$$h\left(\frac{\epsilon''}{1+\epsilon}\right) \leq h\left(\frac{\epsilon}{1+\epsilon}\right). \quad (A22)$$

Thus we can replace ϵ'' with ϵ in the second term as well. This leaves us with the desired expression. ∎

APPENDIX B: MATRIX OPERATIONS IN DISPLACED BASIS

Recall our basis is $\{|i\rangle_A \otimes |n_{\beta_i}\rangle_B\}$. We calculate the matrix elements of certain operators in this basis and evaluate the action of relevant channels.

1. Operators

Our constraint operators take a particularly simple form in the displaced basis. The matrix elements are

$$\langle i|\langle m_{\beta_i}|\left(|k\rangle\langle k| \otimes \hat{n}_{\beta_k}\right)|j\rangle|n_{\beta_j}\rangle = \delta_{ik}\delta_{jk}\langle m_{\beta_k}|\hat{n}_{\beta_k}|n_{\beta_k}\rangle$$

(B1)

$$= \delta_{ik}\delta_{jk}\delta_{mn}n. \quad \text{(B2)}$$

Similarly, for $\hat{n}_{\beta_k}^2$, they are $\delta_{ik}\delta_{jk}\delta_{mn}n^2$. The key map POVMs are more complicated. Recall the POVM elements [Eq. (53)] are

$$P^k = \mathbb{1}_A \otimes R_B^k, \quad \text{(B3)}$$

where R_B^k are the region operators for the nondiscarded signals. The matrix elements are

$$P_{ijmn}^k = \langle i|\langle m_{\beta_i}|\left(\mathbb{1}_A \otimes R_B^k\right)|j\rangle|n_{\beta_j}\rangle \quad \text{(B4)}$$

$$= \delta_{ij}\langle m_{\beta_i}|R_B^k|n_{\beta_j}\rangle \quad \text{(B5)}$$

$$= \langle m_{\beta_i}|R_B^k|n_{\beta_i}\rangle \quad \text{(B6)}$$

$$= \frac{1}{\pi}\int_{\Delta_a}^{\infty}\int_{[(2k-1)\pi/4]+\Delta_p}^{[(2k+1)\pi/4]-\Delta_p} re^{-|\kappa|^2}\frac{\kappa^m\kappa^{*n}}{\sqrt{m!n!}}d\theta\,dr, \quad \text{(B7)}$$

where $\kappa = re^{i\theta} - \beta_i$. This integral is computed in MATLAB.

2. Channels

Our basis for the bipartite Hilbert space is not of the form $|i\rangle_A \otimes |j\rangle_B$, where $|i\rangle_A$ and $|j\rangle_B$ are bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Matrix multiplication proceeds as normal, since we simply have some orthonormal basis. However, operations that care about subsystems, namely the partial trace and its adjoint, have a different matrix representation than the typical presentation. We have

$$\rho_{AB} = \sum_{i,j,m,n} c_{ijmn}|i\rangle\langle j| \otimes |m_{\beta_i}\rangle\langle n_{\beta_j}|, \quad \text{(B8)}$$

where the coefficients c are the matrix elements of ρ. We denote this matrix by M_ρ;

$$M_\rho = \sum_{i,j,m,n} c_{ijmn}|i\rangle\langle j| \otimes |m\rangle\langle n|. \quad \text{(B9)}$$

The reduced density matrix is

$$\rho_A = \sum_{i,j,m,n} c_{ijmn}|i\rangle\langle j|\langle n_{\beta_j}|m_{\beta_i}\rangle. \quad \text{(B10)}$$

Defining

$$G = \sum_{i,j,m,n} \langle n_{\beta_j}|m_{\beta_i}\rangle|i\rangle\langle j| \otimes |m_{\beta_i}\rangle\langle n_{\beta_j}|, \quad \text{(B11)}$$

we have that

$$\langle i|\rho_A|j\rangle = \rho_{ij} \odot G_{ij}, \quad \text{(B12)}$$

where the subscripts on the bipartite operators indicate the respective block matrix, and $\odot$ is the elementwise dot product. Note that each G_{ij} can be thought of as a basis change unitary in $\mathcal{H}_B$. An explicit formula for the elements of G is

$$\langle n_{\beta_j}|m_{\beta_i}\rangle = \langle n|D^\dagger(\beta_j)D(\beta_i)|m\rangle \quad \text{(B13)}$$

$$= \exp[i\,\text{Im}(-\beta_j\beta_i^*)]\langle n|D(\beta_i - \beta_j)|m\rangle \quad \text{(B14)}$$

$$= \exp\left(i\,\text{Im}(-\beta_j\beta_i^*) - \frac{|\beta_i - \beta_j|^2}{2}\right)\sqrt{m!n!}$$

$$\times \sum_{k=0}^{\min(m,n)}\frac{1}{k!(m-k)!(n-k)!}(\beta_i - \beta_j)^{n-k}(\beta_j^* - \beta_i^*)^{m-k}. \quad \text{(B15)}$$

We compute and store this matrix once at the beginning of the optimization algorithm, and use it each time to calculate the partial trace.

The adjoint of the partial trace also has a matrix representation involving G. The adjoint of the partial trace is $\xi(\sigma_A) = \sigma_A \otimes \mathbb{1}_B$. Letting

$$\sigma_A = \sum_{ij} c_{ij}|i\rangle\langle j|, \quad \text{(B16)}$$

we seek d_{ijmn} such that

$$\sigma_A \otimes \mathbb{1}_B = \sum_{ijmn} d_{ijmn}|i\rangle\langle j| \otimes |m_{\beta_i}\rangle\langle n_{\beta_j}|. \quad \text{(B17)}$$

This implies

$$\sum_{mn} d_{ijmn}|m_{\beta_i}\rangle\langle n_{\beta_j}| = c_{ij}\mathbb{1}_B \quad \forall i,j. \quad \text{(B18)}$$

Taking the bra-ket on both sides, we obtain $d_{ijmn} = c_{ij}\langle m_{\beta_i}|n_{\beta_j}\rangle$. We recognize the factor on the right-hand side as G^*. Thus, we have that

$$\xi(\sigma_A) = \sum_{ijmn} c_{ij}|i\rangle\langle j| \otimes G_{ij}^*. \quad \text{(B19)}$$

APPENDIX C: NUMERIC FRAMEWORK MODIFICATION

In principle, the numerical framework presented in Ref. [8] is tight. We observe the following issue in practice. The near-optimal ρ computed in the first step often has constraint violations, due to the inherent imprecision of convex solvers. At all distances and values of excess noise, and for our particular implementation using MATLAB and CVX with the MOSEK solver, these violations are typically $10^{-7} - 10^{-6}$. At distances approaching 200 km, the simulated expectation values $\langle \hat{n}_{\beta_i} \rangle$ and $\langle \hat{n}_{\beta_i}^2 \rangle$ are both small; approximately 10^{-6} for the nonzero values of excess noise we consider. Since the constraint violation is the same order of magnitude as the expectation, the first-step solution is effectively an optimal state for double the excess noise. Given the poor scaling of the protocol with excess noise, this implies the approximate key rate from this first step will be much lower than its theoretical value. Since the first step upper bounds the second step, this implies a poor second-step result. One way to ameliorate this is to solve with a smaller N so that the solver returns a better first-step solution. Thus, purely due to numerical precision issues, solving with $N = 30$ instead of $N = 40$ can improve the key rate at long distances. Even though the correction term $\Delta(W)$ is slightly larger, this is more than offset by the improved quality of the first-step solution. This suggests that due to numerical issues, one should choose the finite dimension carefully, even though analytically a larger dimension is always better.

The reason the first step upper bounds the second step is due to the expansion of the feasible set. Referring to the notation in Appendix D of Ref. [8], the large constraint violations lead to a large value of ϵ', which controls how much the set is expanded for the second step. However, the choice $\epsilon' = \max(\epsilon_{\text{rep}}, \epsilon_{\text{sol}})$ [Eq. (165) of Ref. [8]] is pessimistic. One only needs to choose $\epsilon' = \epsilon_{\text{rep}}$. As noted in Eq. (162) of Ref. [8], this is sufficient to provide a reliable lower bound when accounting for finite numerical precision. Further, note that $f(\rho)$ is lower bounded by a tangent hyperplane at any point in its domain. Thus, it is not necessary to expand the feasible set further to include the point returned by the first step. This change gives improved results in practice, while still being reliable and tight.

In previous work using the numerical framework in Ref. [8], it has been assumed that $\epsilon_{\text{rep}} \leq \epsilon_{\text{sol}}$. Hence the issue of how to suitably choose ϵ_{rep} has not been considered. As noted in Ref. [8], rigorously determining ϵ_{rep} for a particular implementation can be an involved process. For our MATLAB implementation, which has precision better than 10^{-15}, we conservatively use 10^{-10} for ϵ'_{rep} and all elements of $\vec{\epsilon}_{\text{rep}}$. Finally note that in our numerical evaluation of the unbalanced phase-encoded BB84 protocol, we continue to use the original, larger set expansion as in Ref. [8]. This is to ensure a fair comparison to the flag-state squasher numerical results.

APPENDIX D: TRUSTED DETECTOR NOISE OPERATORS

We determine expressions for the relevant operators in the trusted detector noise scenario, focusing on the case where both homodyne detectors in the overall heterodyne setup have the same efficiency η_d and electronic noise ν_{el}. The only change in the protocol for the trusted noise scenario is Bob's POVM. The changed POVM enters the optimization in two different ways: through Bob's new observables in the constraints and new region operators in the objective function definition.

Recall Bob's POVM in the ideal case is a projection onto coherent states $\{1/\pi \, |\zeta\rangle \langle\zeta|\}_{\zeta \in \mathbb{C}}$. In the trusted noise scenario, it is instead a projection onto scaled and displaced thermal states $\left\{(1/\eta_d\pi)\hat{D}\left(\zeta/\sqrt{\eta_d}\right)\rho_{\text{th}}(\bar{n})\hat{D}^\dagger\left(\zeta/\sqrt{\eta_d}\right)\right\}_{\zeta \in \mathbb{C}}$ where the mean photon number of the thermal state is $\bar{n} = (1 - \eta_d + \nu_{\text{el}})/\eta_d$. We simplify our notation by omitting the dependence of ρ_{th} on $\bar{n}$. We denote the POVM elements by G_ζ. Recall our notation $[\cdot]'$ for the noisy version of an operator. From Ref. [39], if an operator in the ideal detector model is defined as

$$X = \int_{\zeta \in \mathbb{C}} f_X(\zeta) \frac{1}{\pi} |\zeta\rangle \langle\zeta| \, d^2\zeta, \tag{D1}$$

then its noisy counterpart is

$$[X]' = \int_{\zeta \in \mathbb{C}} f_X(\zeta) G_\zeta d^2\zeta, \tag{D2}$$

where $d^2\zeta = d\,\text{Re}(\zeta)d\,\text{Im}(\zeta)$.

1. Objective function

Recall the region operators are defined as $R_B^j = 1/\pi \int_{A^j} |\zeta\rangle \langle\zeta| \, d^2\zeta$, where A^j are the regions in phase space in Fig. 2. By definition, the noisy region operators are then

$$\left[R_B^j\right]' = \int_{A^j} G_\zeta \, d^2\zeta, \tag{D3}$$

and the noisy POVM is $\left[P^k\right]' = \mathbb{1}_A \otimes \left[R_B^k\right]'$. Referring to Eq. (B6), the matrix elements are

$$\left[P^k\right]'_{ijmn} = \langle m_{\beta_i} | \left[R_B^k\right]' | n_{\beta_i} \rangle, \tag{D4}$$

$$= \int_{A^k} \langle m_{\beta_i} | G_\zeta | n_{\beta_i} \rangle d^2\zeta, \tag{D5}$$

$$= \frac{1}{\eta_d \pi} \int_{A^k} \langle m | D \left(\frac{\zeta}{\sqrt{\eta_d}} - \beta_i \right) \rho_{\text{th}} D^\dagger \left(\frac{\zeta}{\sqrt{\eta_d}} - \beta_i \right) |n\rangle \, d^2\zeta. \tag{D6}$$

We use the expression in Eq. (B1) in Ref. [39] for the matrix elements of the displaced thermal operator in the Fock basis, convert to polar coordinates, and compute the integral in MATLAB.

2. Observables

To express the observables in the form of Eq. (D1), we write them in antinormal ordering, and replace the ladder operators $\hat{a}, \hat{a}^\dagger$ with ζ, ζ^*,

$$\hat{n} = \hat{a}\hat{a}^\dagger - 1 \implies f_{\hat{n}}(\zeta) = |\zeta|^2 - 1, \tag{D7}$$

$$\hat{n}^2 = \hat{a}^2(\hat{a}^\dagger)^2 - 3\hat{n} - 2 \implies f_{\hat{n}^2}(\zeta) = |\zeta|^4 - 3|\zeta|^2 + 1. \tag{D8}$$

To find f for the displaced observables, we simply perform a change of variables,

$$\hat{n}_\beta = \hat{D}(\beta) \left(\frac{1}{\pi} \int (|\zeta|^2 - 1) |\zeta\rangle \langle \zeta| d^2\zeta \right) \hat{D}^\dagger(\beta), \tag{D9}$$

$$= \frac{1}{\pi} \int (|\zeta|^2 - 1) |\zeta + \beta\rangle \langle \zeta + \beta| d^2\zeta, \tag{D10}$$

$$= \frac{1}{\pi} \int (|\zeta - \beta|^2 - 1) |\zeta\rangle \langle \zeta| d^2\zeta, \tag{D11}$$

and similarly for $\hat{n}_\beta^2$. Thus,

$$f_{\hat{n}_\beta}(\zeta) = |\zeta - \beta|^2 - 1 \tag{D12}$$

and

$$f_{\hat{n}_\beta^2}(\zeta) = |\zeta - \beta|^4 - 3|\zeta - \beta|^2 + 1. \tag{D13}$$

We can now calculate the noisy observables using Eq. (D2). We make use of the following identity:

$$\langle \gamma | \rho_{\text{th}}(\bar{n}) | \gamma \rangle = \frac{e^{-|\gamma|^2/(1+\bar{n})}}{1 + \bar{n}}. \tag{D14}$$

By definition,

$$[\hat{n}_\beta]' = \int (|\zeta - \beta|^2 - 1) G_\zeta \, d^2\zeta \tag{D15}$$

$$= \frac{\eta_d}{\pi} \int \left(\left| \zeta - \frac{\beta}{\sqrt{\eta_d}} \right|^2 - \frac{1}{\eta_d} \right) D(\zeta) \rho_{\text{th}} D^\dagger(\zeta) \, d^2\zeta. \tag{D16}$$

Then,

$$\langle \alpha | [\hat{n}_\beta]' | \alpha \rangle = \frac{\eta_d}{\pi} \int \left(|\zeta - \beta'|^2 - \frac{1}{\eta_d} \right)$$
$$\langle \alpha - \zeta | \rho_{\text{th}} | \alpha - \zeta \rangle \, d^2\zeta \tag{D17}$$

$$= \frac{\eta_d}{\pi(1 + \bar{n})} \int \left(|\zeta + \alpha - \beta'|^2 - \frac{1}{\eta_d} \right) e^{-|\zeta|^2/(1+\bar{n})} d^2\zeta, \tag{D18}$$

where $\beta' = \beta/\sqrt{\eta_d}$. Converting to polar coordinates, the integral is

$$\langle \alpha | [\hat{n}_\beta]' | \alpha \rangle = \frac{\eta_d}{\pi(1 + \bar{n})}$$
$$\times \int \left(r^2 + \gamma^* r e^{i\theta} + \gamma r e^{-i\theta} + |\gamma|^2 - \frac{1}{\eta_d} \right)$$
$$e^{-r^2/(1+\bar{n})} r \, dr \, d\theta \tag{D19}$$

$$= \eta_d |\gamma|^2 + v_{\text{el}}, \tag{D20}$$

where $\gamma = \alpha - \beta'$. By the uniqueness of the Husimi Q-function,

$$[\hat{n}_\beta]' = \eta_d \hat{n}_{\beta/\sqrt{\eta_d}} + v_{\text{el}} \mathbb{1}. \tag{D21}$$

Similarly,

$$[\hat{n}_\beta^2]' = \frac{1}{\pi \eta_d} \int (|\zeta - \beta|^4 - 3|\zeta - \beta|^2 + 1) \hat{D}\left(\frac{\zeta}{\sqrt{\eta_d}} \right) \rho_{\text{th}} \hat{D}^\dagger \left(\frac{\zeta}{\sqrt{\eta_d}} \right) d^2\zeta \tag{D22}$$

$$= \frac{1}{\pi} \int \left(\eta_d^2 |\zeta - \beta'|^4 - 3\eta_d |\zeta - \beta'|^2 + 1 \right) \hat{D}(\zeta) \rho_{\text{th}} \hat{D}^\dagger(\zeta) \, d^2\zeta. \tag{D23}$$

The Q-function is then

$$\langle \alpha | [\hat{n}_\beta^2]' | \alpha \rangle = \frac{1}{\pi(1 + \bar{n})} \int \left(\eta_d^2 |\zeta + \alpha - \beta'|^4 - 3\eta_d |\zeta + \alpha - \beta'|^2 + 1 \right) e^{-|\zeta|^2/(1+\bar{n})} d^2\zeta \tag{D24}$$

$$= \frac{1}{\pi(1+\bar{n})} \int \left[\eta_d^2(r^4 + 4r^2 |\gamma|^2 + |\gamma|^4) - 3\eta_d(r^2 + |\gamma|^2) + 1 \right] e^{-r^2/(1+\bar{n})} \, r \, dr \, d\theta \tag{D25}$$

$$= 2 \left(1 + 2\nu_{\mathrm{el}} + \nu_{\mathrm{el}}^2 + 2\eta_d(1 + \nu_{\mathrm{el}}) |\gamma|^2 - 3\frac{1 + \nu_{\mathrm{el}}}{2} + \eta_d^2 |\gamma|^4 \frac{1}{2} - \frac{3}{2}\eta_d |\gamma|^2 + \frac{1}{2} \right) \tag{D26}$$

$$= \eta_d^2 |\gamma|^4 + \eta_d(4\nu_{\mathrm{el}} + 1) |\gamma|^2 + 2\nu_{\mathrm{el}}^2 + \nu_{\mathrm{el}}. \tag{D27}$$

Again, by the uniqueness of the Husimi Q-function,

$$\left[\hat{n}_\beta^2 \right]' = \eta_d^2 \hat{n}_{\beta/\sqrt{\eta_d}}^2 + \eta_d(4\nu_{\mathrm{el}} + 1 - \eta_d)\hat{n}_{\beta/\sqrt{\eta_d}}$$
$$+ (2\nu_{\mathrm{el}}^2 + \nu_{\mathrm{el}})\mathbb{1}. \tag{D28}$$

APPENDIX E: SIMULATED EXPECTATIONS AND ERROR-CORRECTION COST

We discuss how the coarse-grained expectations can be determined from a heterodyne measurement, and what the expectation values are for the simulation. We focus on the trusted detector noise scenario, as the ideal detector results can be recovered as a special case.

Bob's measurement results determine a probability density $p(\zeta) = \mathrm{Tr}(\rho G_\zeta)$ over the complex plane. In general, given an observable $\Gamma = \int_{\zeta \in \mathbb{C}} f_\Gamma(\zeta) G_\zeta d^2\zeta$, the expectation value is then

$$\mathrm{Tr}(\rho\Gamma) = \int_{\zeta \in \mathbb{C}} f_\Gamma(\zeta) p(\zeta) d^2\zeta. \tag{E1}$$

For the ith conditional state, Bob's coarse-grained observables are $\hat{n}_{\beta_i}$ and $\hat{n}_{\beta_i}^2$. The corresponding functions $f_{\hat{n}_{\beta_i}}$ and $f_{\hat{n}_{\beta_i}^2}$ are given in Eqs. (D12) and (D13). Using his measurement result $p(\zeta)$, Bob can thus compute the integral (E1) to determine the desired expectations for each conditional state.

Note that for the typical quadratures X and P, we have that

$$X = \frac{1}{\sqrt{2}}(\hat{a}^\dagger + \hat{a}) \implies f_X(\zeta) = \sqrt{2}\,\mathrm{Re}(\zeta), \tag{E2}$$

$$P = \frac{i}{\sqrt{2}}(\hat{a}^\dagger - \hat{a}) \implies f_P(\zeta) = \sqrt{2}\,\mathrm{Im}(\zeta). \tag{E3}$$

Thus, by expanding $f_{\hat{n}_{\beta_i}}(\zeta)$ and $f_{\hat{n}_{\beta_i}^2}(\zeta)$ as polynomials in $\mathrm{Re}(\zeta)$ and $\mathrm{Im}(\zeta)$, we can also relate the expectations of $\hat{n}_{\beta_i}$ and $\hat{n}_{\beta_i}^2$ to the moments and crossterms of the measurement data of quadratures X and P.

We now consider the expectations under the simulated channel model. After passing through a Gaussian channel with loss η and excess noise ξ, a coherent signal state

becomes a displaced thermal state

$$|\alpha_i\rangle \langle \alpha_i| \rightarrow \hat{D}(\beta_i) \rho_{\mathrm{th}}\left(\tfrac{\delta}{2}\right) \hat{D}^\dagger(\beta_i), \tag{E4}$$

where $\delta = \eta\xi$ and $\beta_i = \sqrt{\eta}\alpha_i$. The expectation values for each conditional state are straightforward to calculate,

$$\mathrm{Tr}\left[\hat{n}_{\beta_i}\hat{D}(\beta_i) \rho_{\mathrm{th}}\left(\tfrac{\delta}{2}\right) \hat{D}^\dagger(\beta_i) \right] = \mathrm{Tr}\left[\hat{n}\rho_{\mathrm{th}}\left(\tfrac{\delta}{2}\right) \right] = \tfrac{\delta}{2} \tag{E5}$$

and

$$\mathrm{Tr}\left[\hat{n}_{\beta_i}^2 \hat{D}(\beta_i) \rho_{\mathrm{th}}\left(\tfrac{\delta}{2}\right) \hat{D}^\dagger(\beta_i) \right] = \frac{\delta(1 + \delta)}{2}. \tag{E6}$$

For the reduced state constraint, we use the well-known formula for the overlap of two coherent states [48]

$$\langle \alpha_j | \alpha_i \rangle = \exp\left(i\,\mathrm{Im}(\alpha_i\alpha_j^*) - \frac{1}{2}\left| \alpha_i - \alpha_j \right|^2 \right). \tag{E7}$$

The error-correction cost is determined by the simulated joint probability distribution. Given Alice prepares $|\alpha_i\rangle$, the probability Bob gets the key map outcome j, with $i \in \{0, 1, 2, 3\}$ and $j \in \{0, 1, 2, 3, \perp\}$, is given by the following integral:

$$p(j|i) = \mathrm{Tr}\left[R^j \hat{D}(\beta_i) \rho_{\mathrm{th}}\left(\tfrac{\delta}{2}\right) \hat{D}^\dagger(\beta_i) \right] \tag{E8}$$

$$= \int_{A^j} \mathrm{Tr}\left[G_\zeta \hat{D}(\beta_i) \rho_{\mathrm{th}}\left(\tfrac{\delta}{2}\right) \hat{D}^\dagger(\beta_i) \right] d^2\zeta. \tag{E9}$$

The integrand, which is the overlap of two displaced thermal states, is given by [39]

$$\mathrm{Tr}\left[G_\zeta \hat{D}(\beta_i) \rho_{\mathrm{th}}\left(\tfrac{\delta}{2}\right) \hat{D}^\dagger(\beta_i) \right]$$
$$= \frac{1}{\pi(1 + \frac{1}{2}\eta_d\delta + \nu_{\mathrm{el}})} \exp\left(\frac{-\left| \zeta - \sqrt{\eta_d}\beta_i \right|^2}{1 + \frac{1}{2}\eta_d\delta + \nu_{\mathrm{el}}} \right). \tag{E10}$$

The integral in Eq. (E9) is converted to polar coordinates and computed in MATLAB. As the signal states are distributed uniformly, $p_A(i) = \frac{1}{4}$. Then, $p_{AB}(i,j) = \frac{1}{4}p(j|i)$.

Since discarded signals do not incur an error-correction cost, we remove the outcome $j = \perp$ and renormalize p accordingly. Denoting this sifted probability distribution by q, the error-correction cost is

$$\delta_{\text{EC}}^{\text{leak}} = 2 - \beta_{\text{EC}}[H(q_A) + H(q_B) - H(q_{AB})]. \quad (E11)$$

[1] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (IEEE, New York, USA, 1984), p. 175.

[2] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, Phys. Rev. Lett. **67**, 661 (1991).

[3] R. Colbeck and R. Renner, No extension of quantum theory can have improved predictive power, Nat. Commun. **2**, 411 (2011).

[4] 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).

[5] 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).

[6] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Advances in quantum cryptography, Adv. Opt. Photonics **12**, 1012 (2020).

[7] P. J. Coles, E. M. Metodiev, and N. Lütkenhaus, Numerical approach for unstructured quantum key distribution, Nat. Commun. **7**, 11712 (2016).

[8] A. Winick, N. Lütkenhaus, and P. J. Coles, Reliable numerical key rates for quantum key distribution, Quantum **2**, 77 (2018).

[9] N. J. Beaudry, T. Moroder, and N. Lütkenhaus, Squashing Models for Optical Measurements in Quantum Communication, Phys. Rev. Lett. **101**, 093601 (2008).

[10] T. Tsurumaru, Squash operator and symmetry, Phys. Rev. A **81**, 012328 (2010).

[11] O. Gittsovich, N. J. Beaudry, V. Narasimhachar, R. R. Alvarez, T. Moroder, and N. Lütkenhaus, Squashing model for detectors and applications to quantum-key-distribution protocols, Phys. Rev. A **89**, 012325 (2014).

[12] Y. Zhang, P. J. Coles, A. Winick, J. Lin, and N. Lütkenhaus, Security proof of practical quantum key distribution with detection-efficiency mismatch, Phys. Rev. Res. **3**, 013076 (2021).

[13] 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).

[14] T. C. Ralph, Continuous variable quantum cryptography, Phys. Rev. A **61**, 010303 (1999).

[15] M. Hillery, Quantum cryptography with squeezed states, Phys. Rev. A **61**, 022309 (2000).

[16] C. Silberhorn, T. C. Ralph, N. Lütkenhaus, and G. Leuchs, Continuous Variable Quantum Cryptography: Beating the 3 dB Loss Limit, Phys. Rev. Lett. **89**, 167901 (2002).

[17] F. Grosshans and P. Grangier, Continuous Variable Quantum Cryptography Using Coherent States, Phys. Rev. Lett. **88**, 057902 (2002).

[18] F. Grosshans, G. Van Assche, J. Wenger, R. Brouri, N. J. Cerf, and P. Grangier, Quantum key distribution using Gaussian-modulated coherent states, Nature **421**, 238 (2003).

[19] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, Quantum Cryptography Without Switching, Phys. Rev. Lett. **93**, 170504 (2004).

[20] Y.-B. Zhao, M. Heid, J. Rigas, and N. Lütkenhaus, Asymptotic security of binary modulated continuous-variable quantum key distribution under collective attacks, Phys. Rev. A **79**, 012307 (2009).

[21] K. Brádler and C. Weedbrook, Security proof of continuous-variable quantum key distribution using three coherent states, Phys. Rev. A **97**, 022310 (2018).

[22] S. Ghorai, P. Grangier, E. Diamanti, and A. Leverrier, Asymptotic Security of Continuous-Variable Quantum Key Distribution with a Discrete Modulation, Phys. Rev. X **9**, 021059 (2019).

[23] J. Lin, T. Upadhyaya, and N. Lütkenhaus, Asymptotic Security Analysis of Discrete-Modulated Continuous-Variable Quantum Key Distribution, Phys. Rev. X **9**, 041064 (2019).

[24] T. Matsuura, K. Maeda, T. Sasaki, and M. Koashi, Finite-size security of continuous-variable quantum key distribution with digital signal processing, Nat. Commun. **12**, 252 (2021).

[25] N. Killoran and N. Lütkenhaus, Strong quantitative benchmarking of quantum optical devices, Phys. Rev. A **83**, 052320 (2011).

[26] C. H. Bennett, G. Brassard, and N. D. Mermin, Quantum Cryptography Without Bell's Theorem, Phys. Rev. Lett. **68**, 557 (1992).

[27] F. Grosshans, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier, Virtual entanglement and reconciliation protocols for quantum cryptography with continuous variables, Quantum Info. Comput. **3**, 535 (2003).

[28] M. Curty, M. Lewenstein, and N. Lütkenhaus, Entanglement as a Precondition for Secure Quantum Key Distribution, Phys. Rev. Lett. **92**, 217903 (2004).

[29] 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).

[30] 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).

[31] R. Renner and J. I. Cirac, de Finetti Representation Theorem for Infinite-Dimensional Quantum Systems and Applications to Quantum Cryptography, Phys. Rev. Lett. **102**, 110504 (2009).

[32] A. Winter, Coding theorem and strong converse for quantum channels, IEEE Trans. Inf. Theory **45**, 2481 (1999).

[33] A. Winter, Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints, Commun. Math. Phys. **347**, 291 (2016).

[34] A. Uhlmann, The "transition probability" in the state space of a *-algebra, Rep. Math. Phys. **9**, 273 (1976).

[35] C. A. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum-mechanical states, IEEE Trans. Inf. Theory **45**, 1216 (1999).

[36] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, UK, 2018). 1st ed.

[37] E. Kaur, S. Guha, and M. M. Wilde, Asymptotic security of discrete-modulation protocols for continuous-variable quantum key distribution, Phys. Rev. A **103**, 012412 (2021).

[38] M. Heid and N. Lütkenhaus, Efficiency of coherent-state quantum cryptography in the presence of loss: Influence of realistic error correction, Phys. Rev. A **73**, 052316 (2006).

[39] J. Lin and N. Lütkenhaus, Trusted Detector Noise Analysis for Discrete Modulation Schemes of Continuous-Variable Quantum Key Distribution, Phys. Rev. Appl. **14**, 064030 (2020).

[40] M. Grant, S. Boyd, CVX, Matlab Software for Disciplined Convex Programming, version 2.1, http://cvxr.com/cvx (2014).

[41] M. C. Grant and S. P. Boyd, Graph Implementations for Nonsmooth Convex Programs, in *Recent Advances in Learning and Control*, edited by V. D. Blondel, S. P. Boyd, and H. Kimura (Springer London, London, UK, 2008), pp. 95–110.

[42] M. ApS, The MOSEK optimization toolbox for MATLAB Manual. Version 8.0.0.60. (2016), https://www.mosek.com/downloads/8.0.0.60.

[43] S. Fossier, E. Diamanti, T. Debuisschert, R. Tualle-Brouri, and P. Grangier, Improvement of continuous-variable quantum key distribution systems by using optical preamplifiers, J. Phys. B: At., Mol., Opt. Phys. **42**, 114014 (2009).

[44] K. C. Toh, M. J. Todd, and R. H. Tütüncü, SDPT3—A matlab software package for semidefinite programming, version 1.3, Optim. Methods Softw. **11**, 545 (1999).

[45] R. H. Tütüncü, K. C. Toh, and M. J. Todd, Solving semidefinite-quadratic-linear programs using SDPT3, Math. Program. **95**, 189 (2003).

[46] I. George, J. Lin, and N. Lütkenhaus, Numerical calculations of the finite key rate for general quantum key distribution protocols, Phys. Rev. Res. **3**, 013274 (2021).

[47] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Theoretical and Mathematical Physics (Springer-Verlag Berlin Heidelberg, New York, USA, 1993).

[48] C. Gerry and P. Knight, *Introductory Quantum Optics* (Cambridge University Press, Cambridge, UK, 2004).