Reliable numerical key rates for quantum key distribution

Patrick J. Coles, Adam Winick, and Norbert Lütkenhaus
Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, N2L3G1 Waterloo, Ontario, Canada

Introduction.—Quantum key distribution (QKD) will play an important role in quantum-safe cryptography. The main theoretical problem in QKD is to calculate how much secret key can be distributed by a given protocol. A crucial practical issue is that the QKD protocols that are easiest to implement with existing optical technology do not necessarily coincide with the protocols that are easiest to analyze theoretically [1]. In addition, device imperfections and side channels may have a significant effect on the key rate. Addressing these issues requires a robust theoretical method for calculating the key rate. Unfortunately, analytical methods are highly technical, are often limited in scope to particular protocols, and invoke inequalities that introduce looseness into the calculation.

We therefore focus our efforts on numerical methods, which are inherently more robust to both device imperfections and changes in protocol structure. Furthermore, numerics can be made user-friendly, such that the user needs only to define the specifications of the protocol and then the computer performs the key rate calculation. As an example, our group recently released a software package for this purpose [17].

The key rate calculation involves minimizing a convex function over all eavesdropping attacks that are consistent with the experimental data [2–4]. When employing numerics, one issue that arises is the efficiency of this optimization. This issue is particularly important for high-dimensional QKD protocols, or protocols with many signal states, since the relevant optimization involves many parameters. In such cases, the computational time can be very long - sometimes days - so it is crucial to implement a high efficiency algorithm.

Another issue with numerics is reliability. This is more subtle but also more important than the efficiency issue. Due to the inherent paranoia in cryptography, it is natural to ask whether numerically calculated key rates are trustworthy. After all, computers have finite numerical precision. Furthermore, optimization algorithms never truly reach the global optimum, as termination conditions always have some non-zero tolerance. Since QKD is now a serious, real-world technology, key rates must come with a security guarantee, and hand-waving at these numerical issues will not suffice.

In this work, we present a numerical method that solves both the reliability and efficiency issues. Our method provides reliable lower bounds on the key rate with arbitrary tightness for finite-dimensional QKD protocols. Furthermore it is highly efficient and typically returns a key rate within seconds or less on one's personal computer. To illustrate our method, we apply it to three practically interesting scenarios. Namely we consider the Trojan-horse attack [5–7], the BB84 protocol with detector efficiency mismatch [10], and the BB84 protocol with phase-coherent signal states [8, 9]. We improve upon literature key rates in all three cases.

Background.—The well-known asymptotic key rate formula [11] is given by the difference of two information-theoretic quantities associated, respectively, with privacy amplification and error correction. These two terms appear in the following expression for the key rate

$$K = \left(\min_{\rho \in \mathbf{S}} f(\rho) \right) - p_{\text{pass}} \cdot \text{leak}_{\text{obs}}^{\text{EC}} \,. \tag{1}$$

We explain this expression in more detail in the attached manuscript. For now, we note that p_{pass} refers to the probability for passing the post-selection (e.g., sifting) in the protocol, $\text{leak}_{\text{obs}}^{\text{EC}}$ denotes the number of bits of information that Alice publicly reveals during error correction, and $f(\rho)$ is a convex function of the state ρ. Here, ρ is the state shared by Alice, Bob, and possibly other parties involved the protocol. While ρ is unknown, the asymptotic experimental data gives linear constraints on it, of the form

$$\text{Tr}(\Gamma_i \rho) = \gamma_i, \quad \forall i \,, \tag{2}$$

where the Γ_i are Hermitian operators. Let $\mathbf{S}$ denote the set of states that satisfy these constraints

$$\mathbf{S} = \{ \rho \in \mathbf{H}_+ \mid \text{Tr}(\Gamma_i \rho) = \gamma_i, \forall i \} \,, \tag{3}$$

where $\mathbf{H}_+$ is the set of positive semidefinite operators. Also, we add the identity to the set $\{\Gamma_i\}$ to enforce that $\text{Tr}(\rho) = 1$, giving a total of n constraints.

The key rate calculation is an optimization problem, since we must consider the worst-case scenario (the most powerful eavesdropping attack) that is consistent with the experimental data. Hence Eq. (1) involves minimizing over all $\rho \in \mathbf{S}$. Note that the both p_{pass} and $\text{leak}_{\text{obs}}^{\text{EC}}$ are exactly determined by the observations, and hence we can pull them out of the optimization in (1).

Main Result.—We now show how to lower bound the minimization problem in (1). Our strategy is to break up the key rate calculation into two steps:

- Step 1: Find an eavesdropping attack that is close to optimal, which gives an upper bound on the key rate.
- Step 2: Convert this upper bound to a lower bound on the key rate.

With our approach, Step 1 does not need to be perfect - any eavesdropping attack may be used as an input for Step 2. However, if Step 1 returns the optimal attack, our lower bound calculated by Step 2 will be tight. Furthermore, our method for Step 2 is continuous around the optimal attack. Thus, finding a near-optimal attack, produces a near-optimal lower bound.

Step 1 may be solved in various ways using convex optimization methods [12]. For concreteness, the attached manuscript presents one such method, which exploits the structure of our problem and is relatively fast. On the other hand, our main result is a method for performing Step 2, stated in the following theorem.

Theorem 1: Given any $\rho \in \mathbf{S}$, then

$$\left(\min_{\rho\in\mathbf{S}} f(\rho)\right) \geqslant \beta(\rho)\,, \tag{4}$$

where

$$\beta(\sigma) := f(\sigma) - \mathrm{Tr}(\sigma^T \nabla f(\sigma)) + \max_{\vec{y}\in\mathbf{S}^*(\sigma)} \vec{\gamma}\cdot\vec{y}, \tag{5}$$

$$\mathbf{S}^*(\sigma) := \left\{ \vec{y}\in\mathbb{R}^n \mid \sum_i y_i \Gamma_i^T \leqslant \nabla f(\sigma) \right\}. \tag{6}$$

Here, $\vec{\gamma} = \{\gamma_i\}$ is the vector of expectation values from (2). Furthermore, equality in (4) holds if ρ corresponds to an optimal attack.

Figure 1 illustrates the basic idea of Theorem 1. Theorem 1 takes any feasible eavesdropping attack ρ, which gives an upper bound the key rate, and converts it into a reliable lower bound on the key rate. The fact that (5) involves a maximization is crucial for the reliability of the calculation. Since maximization involves approaching the solution from below, every number that the computer outputs is a lower bound on α, even if the computer does not reach the global maximum.

In the attached manuscript, we generalize Theorem 1 to a bound that holds even when one's computer suffers from numerical imprecision, e.g., in storing the variables $\{\Gamma_i\}$ and $\{\gamma_i\}$. Furthermore, we prove that our method yields arbitrarily tight bounds on the key rate, i.e., there is no looseness in our method.

Examples.—In what follows we apply our method to three examples of practical importance.

Trojan-horse attack.—A well-known hacking attack on the phase-encoded BB84 protocol is the Trojan-horse attack [13]. This exploits the fact that Alice's phase modulator is not isolated from Eve. Eve sends a pulse of light,

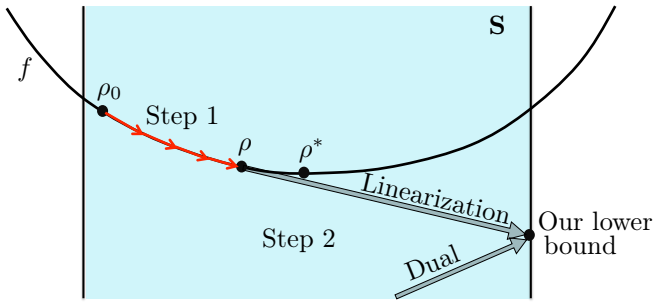

FIG. 1: Illustration of our lower-bounding method. Step 1 is any algorithm that takes an initial feasible point ρ_0 and outputs another feasible point ρ, which may or may not be close to the optimal attack ρ^*. Note that $f(\rho)$ provides an upper bound on $f(\rho^*)$ and, hence, on the key rate. Step 2 converts this into a lower bound, by solving the dual problem of the linearization of f about point ρ. Since the linearization undercuts the curve f and since the dual problem is a maximization, our lower bound is reliable even if the numerical calculation does not reach the global optimum.

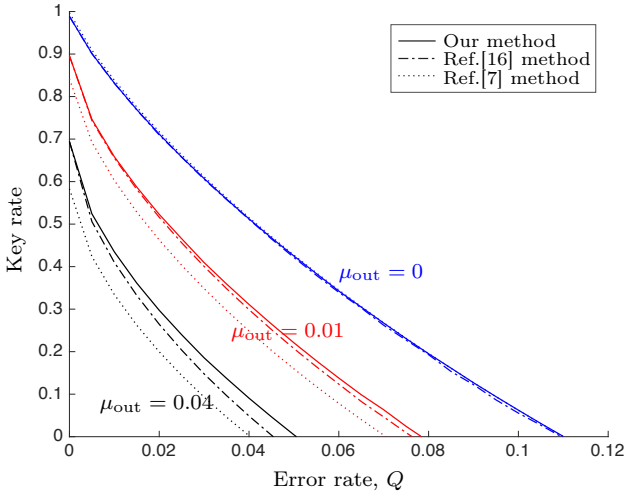

FIG. 2: Key rate vs error rate for the single-photon BB84 protocol under a Trojan-horse attack. The key rate is plotted for different values of μ_{out}. Our numerical method improves on our previous numerical approach in Ref. [16], which in turn gives higher key rates than the analytical method of Ref. [7].

some of which passes through Alice's phase modulator and reflects back to Eve, carrying the information about Alice's encoding. Let μ_{out} be the mean photon number of the light reflected back to Eve. Figure 2 shows the key rate for several values of μ_{out}, for the case where the signal is a single photon. Our method gives higher key rates than those from our previous numerical approach [16], which in turn gives higher key rates than an analytical analysis from Ref. [7].

Efficiency mismatch.—Detector efficiency mismatch is an important issue in QKD because it leads to hacking

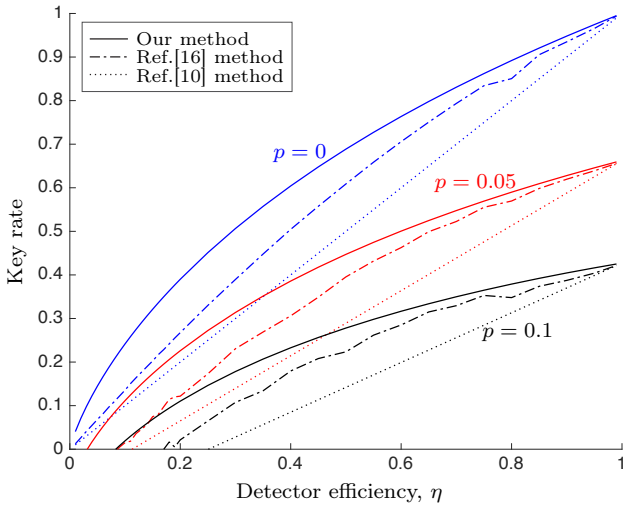

FIG. 3: Key rate for the BB84 protocol with detector efficiency mismatch. Curves are shown for three values of depolarizing probability p (0, 0.05, 0.1). The x-axis is the efficiency of the least efficient detector, with the other detector's efficiency being set to one.

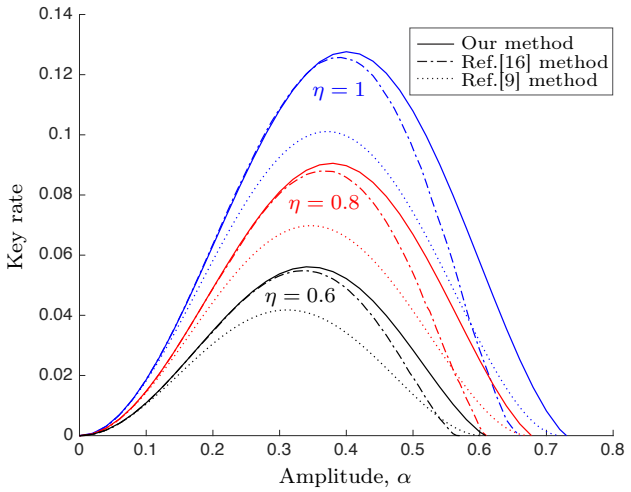

FIG. 4: Key rate versus signal-state amplitude α for three values of transmission probability - $\eta = 1, 0.8, 0.6$ - for the Huttner et al. [8] protocol.

attacks if not accounted for [14, 15]. Ref. [10] gave an analytical lower bound on the key rate in the case of efficiency mismatch, assuming Bob receives at most one photon. For comparison, we assume one detector has perfect efficiency and the other has efficiency η. Figure 3 shows the result of our numerics for this scenario. For all values of η, our method gives higher key rates than the method of Ref. [16], which in turn are higher than those of Ref. [10].

BB84 with phase-coherent signals.—Finally, consider a protocol proposed by Huttner et al. [8] and analyzed by Lo and Preskill [9]. This is a phase-encoded BB84 protocol, but using coherent states of amplitude α instead of single-photon states. This is quite practical because the experimenter does not need to do phase randomization. Ref. [9] gave an analytical lower bound on the key rate for this protocol, as a function of transmission probability η and amplitude α. Their theoretical curves are shown as dotted lines in Fig. 4, for several values of η. In the same plot, we show the result of our numerical optimization as solid lines, with the key rates obtained from the method in Ref. [16] shown as dashed-dotted lines. Interestingly our numerics give higher key rates than the previous literature over the entire parameter range.

Conclusions.—Reliability is the most important issue with numerical key rate calculations, since key rates must come with a security guarantee. In this work, we presented a method that solves the reliability issue, while retaining the efficiency of convex optimization and eliminating all looseness from the calculation.

[1] V. Scarani, H. Bechmann-Pasquinucci, N. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Reviews of Modern Physics **81**, 1301 (2009).

[2] R. Renner, Ph.D. thesis, ETH Zurich (2005).

[3] R. Renner, N. Gisin, and B. Kraus, Physical Review A **72**, 012332 (2005), ISSN 1050-2947.

[4] S. Watanabe, R. Matsumoto, and T. Uyematsu, Physical Review A **78**, 042316 (2008), ISSN 10502947.

[5] A. Vakhitov, V. Makarov, and D. R. Hjelme, Journal of Modern Optics **48**, 2023 (2001).

[6] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, Physical Review A **73**, 022320 (2006), ISSN 1050-2947.

[7] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. L. Yuan, and A. J. Shields, Physical Review X **5**, 1 (2015).

[8] B. Huttner, N. Imoto, N. Gisin, and T. Mor, Physical Review A **51**, 1863 (1995), 9502020.

[9] H. Lo and J. Preskill, Quantum Information and Computation **7**, 431 (2007).

[10] C. C.-H. F. Fung, K. Tamaki, B. Qi, H.-K. H. Lo, and X. Ma, Quantum Inf.˜Comput. **9**, 0131 (2009), 0802.3788.

[11] I. Devetak and A. Winter, Proceedings of the Royal Society A **461**, 207 (2005).

[12] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).

[13] N. Jain, B. Stiller, I. Khan, V. Makarov, C. Marquardt, and G. Leuchs, IEEE Journal on Selected Topics in Quantum Electronics **21** (2015).

[14] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Nature Photonics **4**, 5 (2010).

[15] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Nature Communications **2**, 349 (2011).

[16] P. J. Coles, E. M. Metodiev, and N. Lütkenhaus, Nature Communications **7**, 11712 (2016), ISSN 2041-1723, 1510.01294.

[17] This software can be downloaded from the website: https://lutkenhausgroup.wordpress.com/qkd-software/.

Reliable numerical key rates for quantum key distribution

Patrick J. Coles, Adam Winick, and Norbert Lütkenhaus
Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, N2L3G1 Waterloo, Ontario, Canada

The holy grail of quantum key distribution (QKD) theory is a robust, quantitative method to explore novel protocol ideas and to investigate the effects of device imperfections on the key rate. We argue that numerical methods are superior to analytical ones for this purpose. However, new challenges arise with numerical approaches, including the efficiency (i.e., possibly long computation times) and reliability of the calculation. In this work, we present a reliable, efficient, and tight numerical method for calculating key rates for finite-dimensional QKD protocols. We illustrate our approach by finding higher key rates than those previously reported in the literature for several interesting scenarios (e.g., the Trojan-horse attack and the phase-coherent BB84 protocol). Our method will ultimately improve our ability to automate key rate calculations and, hence, to develop a user-friendly software package that could be used widely by QKD researchers.

Contents

I. INTRODUCTION

The possibility of large-scale quantum computers in the near future has spawned the field of quantum-safe cryptography [1]. This includes classical techniques based on computational hardness as well as information-theoretic approaches via physical assumptions. The latter invokes either assumptions about the physical channel [2], or less restrictive, the assumption only that the laws of quantum physics are correct, which is the basis for quantum key distribution (QKD). See Ref. [3] for a review of QKD and Ref. [4] for an update of recent progress. Both classical methods and QKD will likely play a role in quantum-safe cryptographic implementations.

The maturity of QKD technology is evidenced by a recent QKD satellite launch [5] as well as developments of fiber-based networks [6–8], suggesting that global networks are on the horizon. Still, there remains important open problems in QKD theory, such as (1) optimizing the practicality of QKD protocols to make them easily implementable, and (2) understanding the effects of device imperfections and side channels.

Solving these problems requires a robust theoretical method for evaluating a QKD protocol's performance, which involves a detailed security analysis. Performance

is then quantified by the key rate - the number of bits of secret key obtained per exchange of quantum signal. Unfortunately, analytical methods for calculating the key rate are highly technical, are often limited in scope to particular protocols, and invoke inequalities that introduce looseness into the calculation.

We therefore focus our efforts on numerical methods, which are inherently more robust to both device imperfections and changes in protocol structure. Furthermore, numerics can be made user-friendly, such that the user needs only to define the specifications of the protocol and then the computer performs the key rate calculation. As an example, our group recently released a software package for this purpose.[1]

The key rate calculation involves minimizing a convex function over all eavesdropping attacks that are consistent with the experimental data [9–11]. When employing numerics, one issue that arises is the efficiency of this optimization. This issue is particularly important for high-dimensional QKD protocols, or protocols with many signal states, since the relevant optimization involves many parameters. In such cases, the computational time can be very long - sometimes days - so it is crucial to implement a high efficiency algorithm.

Another issue with numerics is reliability. This is more subtle but also more important than the efficiency issue. Due to the inherent paranoia in cryptography, it is natural to ask whether numerically calculated key rates are trustworthy. After all, computers have finite numerical precision. Furthermore, optimization algorithms never truly reach the global optimum, as termination conditions always have some non-zero tolerance. Since QKD is now a serious, real-world technology, key rates must come with a security guarantee, and hand-waving at these numerical issues will not suffice.

In this work, we present a numerical method that solves both the reliability and efficiency issues. Our method provides reliable lower bounds on the key rate with arbitrary tightness for finite-dimensional QKD protocols. Furthermore it is highly efficient and typically returns a key rate within seconds or less on one's personal computer. To illustrate our method, we apply it to three practically interesting scenarios. Namely we consider the trojan horse attack [12–14], the BB84 protocol with phase-coherent signal states [15, 16], and the BB84 protocol with detector efficiency mismatch [17]. We improve upon literature key rates in all three cases.

Directly calculating the key rate involves a minimization problem (see Sec. II). When solving this on a computer, the algorithm will terminate before reaching the global optimum and hence will return an upper bound on the true key rate. However, we are interested in reliable lower bounds, i.e., achievable key rates. In previous work [18], we noted this issue as motivation for transforming the optimization problem to the so-called dual problem [19]. This transforms the minimization problem into a maximization problem. Therefore, the dual problem will return a lower bound on the key rate, as desired. This method led to novel insights for particular protocols as discussed in [18]. However, in order to simplify the optimization in the dual problem, we invoked an inequality that in some cases introduces looseness into the key rate and, furthermore, makes the optimization problem non-convex. Ultimately the non-convexity reduces the efficiency of this approach, making it difficult to apply to protocols with large numbers of signal states.

We therefore present a method here that retains the efficiency of convex optimization, but also has the reliability of the dual problem. Our approach is to break up the calculation into two steps. The first step approximately minimizes the convex function, and hence finds an eavesdropping attack that is close to optimal. The second step takes this approximately optimal attack and converts it into a lower bound on the key rate. Breaking it up into these two steps adds flexibility to our method, in that any algorithm can be employed for the initial minimization of the convex function.

Our main technical result is to provide a recipe for performing the second step, i.e., for converting a near-optimal attack into a tight lower bound on the key rate. At the technical level, we derive our main result first by linearizing the problem and then by transforming to the dual problem of the subsequent linearized problem. The idea is that, for a convex function, any linearization (about any point) will undercut (and hence lower bound) the curve. One obtains the tightest lower bound by this method if one linearizes about a point corresponding to the global minimum of the convex function.

In what follows we first give background on key rate calculations in the next section. Then we present our main result in Sec. III. In Sec. IV we describe how our approach applies to a general class of QKD protocols. We illustrate our method for three interesting example protocols in Sec. V, and finally we conclude in Sec. VI. Technical details can be found in the Appendix.

II. BACKGROUND

The well-known asymptotic key rate formula [20] is given by the difference of two information-theoretic quantities associated, respectively, with privacy amplification (PA) and error correction (EC). These two terms appear in the following expression for the key rate

$$K = p_{\text{pass}} \left(\min_{\rho \in \mathbf{S}} \widehat{f}(\rho) - \text{leak}_{\text{obs}}^{\text{EC}} \right) \tag{1}$$

$$= \left(\min_{\rho \in \mathbf{S}} f(\rho) \right) - p_{\text{pass}} \cdot \text{leak}_{\text{obs}}^{\text{EC}} . \tag{2}$$

We explain this expression in more detail in Sec. IV. For now, we note that p_{pass} refers to the probability

[1] This software can be downloaded from the website: https://lutkenhausgroup.wordpress.com/qkd-software/.

for passing the post-selection (e.g., sifting) in the protocol, $\text{leak}_{\text{obs}}^{\text{EC}}$ denotes the number of bits of information that Alice publicly reveals during error correction, and $f(\rho) = p_{\text{pass}} \cdot \widehat{f}(\rho)$ is a function defined below in Eq. (5).

The first term in (1) is the PA term. Here, ρ is the density operator shared by Alice, Bob, and possibly other parties involved the protocol. (Note that prepare-and-measure protocols can be recast as entanglement-based protocols and are described by same mathematics, see Sec. IV for elaboration.) This density operator ρ is unknown, but the asymptotic experimental data gives linear constraints on it, of the form

$$\text{Tr}(\Gamma_i \rho) = \gamma_i, \quad \forall i\,, \tag{3}$$

where the Γ_i are Hermitian operators. Let $\mathbf{S}$ denote the set of states that satisfy these constraints

$$\mathbf{S} = \{\rho \in \mathbf{H}_+ \mid \text{Tr}(\Gamma_i \rho) = \gamma_i, \forall i\}\,, \tag{4}$$

where $\mathbf{H}_+$ is the set of positive semidefinite operators. Also, we add the identity to the set $\{\Gamma_i\}$ to enforce that $\text{Tr}(\rho) = 1$, giving a total of n constraints.

The key rate calculation is an optimization problem, since we must consider the worst-case scenario (the most powerful eavesdropping attack) that is consistent with the experimental data. Hence Eq. (1) involves minimizing over all $\rho \in \mathbf{S}$. Note that the error correction term is exactly determined by the observations, and hence we can pull it out of the optimization in (2). Similarly, p_{pass} is known from the observations and is pulled into the optimization in (2).

As discussed in Sec. IV, $f(\rho)$ can be written as

$$f(\rho) = D(\mathcal{G}(\rho)||\mathcal{Z}(\mathcal{G}(\rho))\,, \tag{5}$$

where $D(\sigma||\tau) := \text{Tr}(\sigma \log \sigma) - \text{Tr}(\sigma \log \tau)$ is the relative entropy, $\mathcal{G}$ is a completely positive (CP) map, and $\mathcal{Z}$ is a completely positive trace preserving (CPTP) map, more specifically a pinching quantum channel. (Sec. IV discusses the meaning of $\mathcal{G}$ and $\mathcal{Z}$, which are respectively related to the post-selection and the key map of the QKD protocol). Due to the joint convexity of the relative entropy, the function $f(\rho)$ is convex in ρ. Furthermore the problem

$$\alpha := \min_{\rho \in \mathbf{S}} f(\rho) \tag{6}$$

is a convex optimization problem since the set $\mathbf{S}$ is convex (see, e.g., Ref. [19]). While efficient numerical methods are known for such convex problems, the key rate calculation is unique compared to other convex problems, in that getting "close" to the optimal point is not good enough. One needs a reliable lower bound on the key rate, i.e., guaranteed security.

III. MAIN RESULT

A. Reliable lower bound

We now show how to lower bound the minimization problem in (6). Our strategy is to break up the key rate calculation into two steps:

- Step 1: Find an eavesdropping attack that is close to optimal, which gives an upper bound on the key rate.
- Step 2: Convert this upper bound to a lower bound on the key rate.

With our approach, Step 1 does not need to be perfect - any eavesdropping attack may be used as an input for Step 2. However, if Step 1 returns the optimal attack, our lower bound calculated by Step 2 will be tight. Furthermore, our method for Step 2 is continuous around the optimal attack. Thus, finding a near-optimal attack, produces a near-optimal lower bound.

Step 1 may be solved in various ways using convex optimization methods [19]. For concreteness, Sec. III E presents one such method, which exploits the structure of our problem and is relatively fast.

On the other hand, our main result is a method for performing Step 2. We approach Step 2 via a sequence of theorems that successively improve the reliability and robustness of the lower bounds which they return. First, Theorem 1 presents the conceptual foundation for our lower bounding method. However, this theorem is stated under a restrictive assumption that is not generally true. Therefore, we extend our result in Theorem 3. Finally, we improve our result once more in Theorem 4, which addresses numerical imprecision and is directly useful for numerical key rate calculations. Sec. III D presents the argument that our method yields arbitrarily tight bounds on the key rate.

We now present our main result in its simplest conceptual form. To state this result, we first define the gradient of f at point ρ, whose representation in the standard basis $\{|j\rangle\}$ is

$$\nabla f(\rho) := \sum_{j,k} d_{jk} |j\rangle\langle k|, \quad \text{with } d_{jk} := \left.\frac{\partial f(\sigma)}{\partial \sigma_{jk}}\right|_{\sigma=\rho} \tag{7}$$

and $\sigma_{jk} := \langle j|\sigma|k\rangle$.

Theorem 1: Given any $\rho \in \mathbf{S}$, if $\nabla f(\rho)$ exists, then

$$\alpha \geqslant \beta(\rho)\,, \tag{8}$$

where α was defined in (6) and

$$\beta(\sigma) := f(\sigma) - \text{Tr}(\sigma^T \nabla f(\sigma)) + \max_{\vec{y} \in \mathbf{S}^*(\sigma)} \vec{\gamma} \cdot \vec{y}\,, \tag{9}$$

$$\mathbf{S}^*(\sigma) := \left\{ \vec{y} \in \mathbb{R}^n \mid \sum_i y_i \Gamma_i^T \leqslant \nabla f(\sigma) \right\}. \tag{10}$$

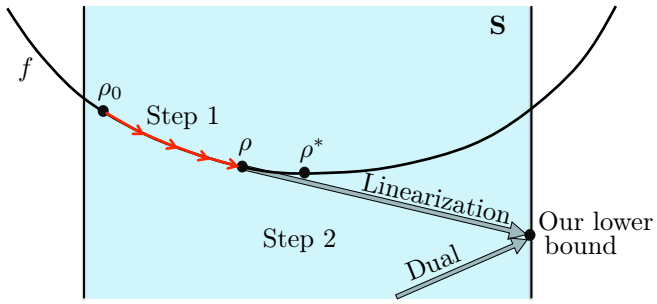

FIG. 1: Illustration of our lower-bounding method. Step 1 is any algorithm that takes an initial feasible point ρ_0 and outputs another feasible point ρ, which may or may not be close to the optimal attack ρ^*. Note that $f(\rho)$ provides an upper bound on $f(\rho^*)$ and, hence, on the key rate. Step 2 converts this into a lower bound, by solving the dual problem of the linearization of f about point ρ. Since the linearization undercuts the curve f and since the dual problem is a maximization, our lower bound is reliable even if the numerical calculation does not reach the global optimum.

Here, the transpose ^{T} is taken in the same basis as that used to define the gradient in (7), and $\vec{\gamma} = \{\gamma_i\}$ is the vector of expectation values from (3). Furthermore, equality in (8) holds if $f(\rho) = \alpha$, i.e., if ρ corresponds to an optimal attack.

The proof is given in Appendix A. It involves linearizing the convex function f about point ρ and then transforming the subsequent linearized problem to its dual problem (see [19] for discussion of duality).

Figure 1 illustrates the basic idea of Theorem 1. Theorem 1 takes any feasible eavesdropping attack ρ, which gives an upper bound $f(\rho)$ on α, and converts it into a reliable lower bound on α. The fact that (9) involves a maximization is crucial for the reliability of the calculation. Since maximization involves approaching the solution from below, every number that the computer outputs is a lower bound on α, even if the computer does not reach the global maximum.

B. A more robust bound

The assumption that ∇f exists over the entirety of $\mathbf{S}$ does not necessarily hold. In particular, the gradient has the form

$$[\nabla f(\rho)]^T = \mathcal{G}^\dagger(\log \mathcal{G}(\rho)) - \mathcal{G}^\dagger(\log \mathcal{Z}(\mathcal{G}(\rho)))\,. \tag{11}$$

By inspection, we see that if $\mathcal{G}(\rho)$ is singular, $\nabla f(\rho)$ may not exist, and hence Theorem 1 would not apply. Ideally we would like a bound that holds for all $\rho \in \mathbf{S}$. Towards this end, it is helpful to first state the following lemma, proved in Appendix B.

Lemma 2: The gradient $\nabla f(\rho)$, defined in (7), exists as long as $\rho > 0$.

The above lemma motivates the following theorem, which is a more robust version of Theorem 1. To state this theorem, it helps to define the notation

$$\rho(\epsilon) := \rho + \epsilon \mathbb{1} \tag{12}$$

for an arbitrary matrix ρ.

Theorem 3: Given any $\rho \in \mathbf{S}$, where ρ is $d \times d$, and $\epsilon \in \mathbb{R}$, define $\rho(\epsilon)$ according to (12). If $0 < \epsilon < 1/(de)$, where e is the base of the natural logarithm, then

$$\alpha \geqslant \beta_\epsilon(\rho(\epsilon)) - \delta_\epsilon \tag{13}$$

where

$$\beta_\epsilon(\sigma) := \beta(\sigma) + \epsilon \mathrm{Tr}(\nabla f(\sigma))\,, \tag{14}$$

$$\delta_\epsilon := -2d\epsilon \log_2 \epsilon\,. \tag{15}$$

The proof is given in Appendix B 4. The basic idea of Theorem 3 is that it is essentially just Theorem 1, but applied to a slightly perturbed state $\rho(\epsilon) > 0$. Note that this theorem generalizes Theorem 1, since we have that

$$\lim_{\epsilon \to 0+} \beta_\epsilon(\rho(\epsilon)) - \delta_\epsilon = \beta(\rho)\,, \tag{16}$$

assuming that $\nabla f(\rho)$ exists. However, Theorem 3 is more robust than Theorem 1, because it holds for any $\rho \in \mathbf{S}$, even if $\nabla f(\rho)$ does not exist.

C. Finite precision computation

Finding a lower bound with our method requires a computer program for any nontrivial QKD protocol. The finite precision inherent to computational methods introduces errors that threaten the reliability of the calculated lower bounds.

With this threat in mind, we identify all possible sources of errors in evaluating Theorem 1. First, the variables will not be precise: ρ will not exactly satisfy the constraints, and $\{\Gamma_i\}$ and $\{\gamma_i\}$ will not be exact. Second, function evaluations will not be precise: every function evaluation introduces errors.

The aforementioned errors fall into broader categories. Variable imprecision is implementation independent since we can characterize the imprecision in a universal manner. Function-evaluation imprecision is implementation dependent since in general, its characterization varies widely with the particular algorithm (particularly for evaluating nontrivial functions such as the matrix logarithm appearing in our objective function). Since the latter kind of errors depend on the implementation, a universal treatment of them is not possible. In principle, it is possible to bound the effect of such errors for a

particular implementation [21], although it is beyond the scope of this article. On the other hand, we give a full treatment of implementation-independent errors in what follows.

From a computational perspective, it is virtually impossible to find an element strictly in $\mathbf{S}$. Furthermore, for many problems, the computer representation $\{\tilde{\Gamma}_i\}$ and $\{\tilde{\gamma}_i\}$ do not equal $\{\Gamma_i\}$ and $\{\gamma_i\}$ by the nature of their numerical construction. This lack of strictly constraint-satisfying density matrices motivates the need for a relaxed theorem.

We show (See Appendix C 1) that both strict set membership and imprecise constraints can be described by the inequalities

$$|\mathrm{Tr}(\tilde{\Gamma}_i\rho) - \tilde{\gamma}_i| \leqslant \epsilon' \,, \forall i \,, \tag{17}$$

where $\epsilon' > 0$. Determining ϵ' depends heavily on how $\{\tilde{\Gamma}_i\}$ and $\{\tilde{\gamma}_i\}$ are constructed, so we leave out a precise analysis. In general, given a suitable high-precision computing environment, ϵ' may be made arbitrarily small. (For example, for our calculations in Sec. V, we choose $\epsilon' < 10^{-12}$.) The bound on the unknown constraint violations motivates the introduction of a relaxed set

$$\mathbf{S}_{\epsilon'} := \left\{ \rho \in \mathbf{H}_+ \middle| |\mathrm{Tr}(\tilde{\Gamma}_i\rho) - \tilde{\gamma}_i| \leqslant \epsilon', \forall i \right\} . \tag{18}$$

So long as ϵ' is larger than the constraint violations, the relation $\mathbf{S} \subseteq \mathbf{S}_{\epsilon'}$ should hold (See Appendix C 1). With this new set, we now present a relaxed version of Theorem 3, with the proof given in Appendix C 3.

Theorem 4: Given any $\rho_{\epsilon'} \in \mathbf{S}_{\epsilon'}$, where $\rho_{\epsilon'}$ is $d \times d$, $\epsilon' > 0$, and $0 < \epsilon < 1/(de)$, define $\rho_{\epsilon'}(\epsilon) := \rho_{\epsilon'} + \epsilon\mathbb{1}$. Then

$$\alpha \geqslant \beta_{\epsilon\epsilon'}(\rho_{\epsilon'}(\epsilon)) - \delta_\epsilon \tag{19}$$

where δ_ϵ was defined in (15) and

$$\beta_{\epsilon\epsilon'}(\sigma) := L_\epsilon(\sigma) + M_{\epsilon'}(\sigma) \,, \tag{20}$$

$$L_\epsilon(\sigma) := f(\sigma) - \mathrm{Tr}(\sigma^T \nabla f(\sigma)) + \epsilon \mathrm{Tr}(\nabla f(\sigma)) \,, \tag{21}$$

$$M_{\epsilon'}(\sigma) := \max_{\vec{y} \in \mathbf{S}^*_{\epsilon'}(\sigma)} \left[(\vec{\tilde{\gamma}}^T + \epsilon', -\vec{\tilde{\gamma}}^T + \epsilon')^T \cdot \vec{y} \right] , \tag{22}$$

$$\mathbf{S}^*_{\epsilon'}(\sigma) := \left\{ \vec{y} \in \mathbb{R}^{2n} \mid \sum_{i=1}^{n} y_i (\tilde{\Gamma}_i^+)^T + \sum_{i=1}^{n} y_{i+n} (\tilde{\Gamma}_i^-)^T \leqslant \nabla f(\sigma) \right\} , \tag{23}$$

$$\tilde{\Gamma}_i^+ := \mathrm{diag}(\tilde{\Gamma}_i, \delta_{i1}, \delta_{i2}, ..., \delta_{in}, \vec{0}^T) \,, \tag{24}$$

$$\tilde{\Gamma}_i^- := \mathrm{diag}(-\tilde{\Gamma}_i, \vec{0}^T, \delta_{i1}, \delta_{i2}, ..., \delta_{in}) \,. \tag{25}$$

Here, δ_{ij} denotes the Kronecker delta and diag denotes the block diagonalization of the set of matrices.

By noting that

$$\lim_{\epsilon' \to 0+} \beta_{\epsilon\epsilon'}(\sigma) = \beta_\epsilon(\sigma), \tag{26}$$

one can see that Theorem 4 generalizes Theorem 3.

The idea is that Theorem 4 provides a method that is robust to constraint violation due to numerical imprecision. Hence, Theorem 4 is directly useful for numerical key rate calculations. Although it is more complicated than Theorems 1 and 3, Theorem 4 is what we employ in practice for our key rate calculations. For example, the calculations presented in Sec. V use this theorem.

D. Convergence and tightness

Theorem 4 is directly useful in key rate calculations, and as such we discuss some of its convergence properties.

In practice, Step 1 does not return a density matrix ρ^* that minimizes f over $\mathbf{S}$. Instead it finds a matrix $\rho_{\epsilon'} \in \mathbf{S}_{\epsilon'}$ that approximately minimizes f over $\mathbf{S}_{\epsilon'}$. Hence, we answer the natural question of how close the lower bound produced from $\rho_{\epsilon'}$ will be to $\alpha = f(\rho^*)$.

Note that by introducing the positive definite state $\rho_{\epsilon'}(\epsilon) = \rho_{\epsilon'} + \epsilon\mathbb{1}$, it follows from Lemma 2 that the lower bound produced by Theorem 4 will be continous with respect to its argument. Thus, if $\rho^*_{\epsilon'}$ minimizes f over $\mathbf{S}_{\epsilon'}$ and $\rho_{\epsilon'}$ is close to $\rho^*_{\epsilon'}$ under some norm, it follows that the lower bounds produced by applying Theorem 4 to $\rho_{\epsilon'}$ and $\rho^*_{\epsilon'}$ will be close. Concretely, we have

$$\lim_{\rho_{\epsilon'} \to \rho^*_{\epsilon'}} \beta_{\epsilon\epsilon'}(\rho_{\epsilon'}(\epsilon)) - \delta_\epsilon = \beta_{\epsilon\epsilon'}(\rho^*_{\epsilon'}(\epsilon)) - \delta_\epsilon \,, \tag{27}$$

showing that our lower bounding method converges (i.e. the optimal lower bound of $\rho^*_{\epsilon'}$ is approachable).

In Appendix D we show that the minimizer ρ^* of f over $\mathbf{S}$ and the minimizer $\rho^*_{\epsilon'}$ of f over $\mathbf{S}_{\epsilon'}$ satisfy

$$\lim_{\epsilon, \epsilon' \to 0} \beta_{\epsilon\epsilon'}(\rho^*_{\epsilon'}(\epsilon)) - \delta_\epsilon = f(\rho^*) \,. \tag{28}$$

Therefore, the lower bound produced by Theorem 4 is tight when applied to $\rho^*_{\epsilon'}$.

Combining (27) and (28), it follows that the lower bound produced by applying Theorem (4) to $\rho_{\epsilon'}$ will be arbitrarily close to $f(\rho^*)$ provided that ϵ, ϵ' are small and $\rho_{\epsilon'}$ is close to $\rho^*_{\epsilon'}$. So provided a suitable computer implementation, our method will produce lower bounds on the true key rate that are arbitrarily tight.

E. Finding a near-optimal attack

Thusfar we focused on Step 2 of the key rate calculation procedure. However, Step 1 needs to be addressed since obtaining a reasonable lower bound from Step 2 requires a $\rho \in \mathbf{S}$ that is sufficiently close to the optimal solution ρ^*. Here we present an algorithm that has proved effective in practice.

Note that in what follows we do not distinguish between exact and inexact representations of $\{\Gamma_i\}$ or $\{\gamma_i\}$, nor do we distinguish between exact and approximate set membership. The validity of this approach is justified by

observing that Step 1 is decoupled from Step 2. Specifically, we can apply Theorem 4 to any positive semidefinite matrix that is approximately consistent with the constraints, and still obtain a reliable lower bound.

By applying the Gram-Schmidt process to the original set of observables $\{\Gamma_i\}$, we obtain a set $\{\bar{\Gamma}_i\}$ of hermitian operators that are orthogonal under the Hilbert-Schmidt norm. The expectation value of each of these operators is denoted

$$\bar{\gamma}_i := \langle \bar{\Gamma}_i \rangle \,. \tag{29}$$

We extend this set to an orthonormal basis $\{\bar{\Gamma}_i\} \cup \{\Omega_j\}$ for the hermitian operator space. With this basis, we may rewrite (4) as

$$\mathbf{S} = \left\{ \sum_i \bar{\gamma}_i \bar{\Gamma}_i + \sum_j \omega_j \Omega_j \in \mathbf{H}_+ \mid \vec{\omega} \in \mathbb{R}^m \right\} , \tag{30}$$

where m is the number of free parameters. This perspective on $\mathbf{S}$ divides the operator space into "fixed" and "free" subspaces. From a practical optimization point-of-view, the benefit of this representation is that it reduces the original equality constrained problem (which is cumbersome to work with) to a constrained minimization subject to a single semidefinite constraint.

We now adapt the Frank-Wolfe method [22] to problem (6).

Algorithm 1 Minimization algorithm for Step 1

1: Let $\epsilon > 0$, $\rho_0 \in \mathbf{S}$ and set $i = 0$.
2: Compute $\Delta\rho := \arg\min_{\Delta\rho} \mathrm{Tr}\left[(\Delta\rho)^T \nabla f(\rho_i)\right]$ subject to $\Delta\rho + \rho_i \in \mathbf{S}$.
3: If $\mathrm{Tr}\left[(\Delta\rho)^T \nabla f(\rho_i)\right] < \epsilon$ then STOP.
4: Find $\lambda \in (0,1)$ that minimizes $f(\rho_i + \lambda\Delta\rho)$.
5: Set $\rho_{i+1} = \rho_i + \lambda\Delta\rho$, $i \leftarrow i+1$ and go to 2.

There are two concrete reasons why adopting the subspace perspective, as in (30), is useful for solving Algorithm 1. First, finding a $\rho_0 \in \mathbf{S}$ becomes simple. We only need to find $\vec{\omega}$ so that

$$\rho_0 = \sum_i \bar{\gamma}_i \bar{\Gamma}_i + \sum_j \omega_j \Omega_j \in \mathbf{H}_+ \,. \tag{31}$$

Second, $\Delta\rho$ has the form

$$\Delta\rho = \sum_j \omega_j \Omega_j \,. \tag{32}$$

Calculating $\vec{\omega}$ requires solving the standard linear semidefinite program (SDP)

$$\vec{\omega} = \arg\min_{\vec{\omega}} \sum_j \omega_j \mathrm{Tr}\left[\Omega_j^T \nabla f(\rho_i)\right] \tag{33}$$

$$\text{subject to} \sum_j (\omega_j \Omega_j) + \rho_i \in \mathbf{H}_+ \,. \tag{34}$$

Problems of this form have been extensively studied (e.g., see [19]) and efficient SDP solvers are widely available.

IV. GENERAL FRAMEWORK FOR QKD PROTOCOLS

1. *Prototypical protocol*

Having stated our main result in abstract form, we now connect our result to concrete QKD protocols. Our goal is to present a framework that applies to the known discrete-variable (DV) QKD protocols in the literature. This includes both entanglement-based (EB) and prepare-and-measure (PM) protocols. Examples of protocols that fall under our framework include the BB84 [23], B92 [24], SARG [3, 25], six-state [26], and decoy-state protocols [27]. Rather than show how our approach applies to each of these examples, we instead construct a generic protocol that encompasses these examples. We call this the "prototypical protocol", shown in Fig. 2. (In Fig. 2, the distinction between the EB and PM scenarios is depicted by a box around the source with a dashed outline, indicating that the source may or may not be located inside Alice's lab.) We will now show how to apply our approach to this prototypical protocol.

We remark that our framework can be easily extended to cover protocols with a central node between Alice and Bob, such as the MDI (measurement-device independent) [28] and the simplified trusted node [29] protocols. We direct the reader to Ref. [18] for a discussion of this extension.

Let us now describe the basic steps involved in our prototypical protocol:

1. In the EB scenario, Alice and Bob each receive a quantum signal (A and B, respectively) from a source and they measure the signal according to the respective POVMs $P^A = \{P_j^A\}$ and $P^B = \{P_j^B\}$, producing the raw data. In the PM scenario, the same mathematical description applies (via the so-called source-replacement scheme [30, 31]) since one can think of Alice's prepared states as resulting from Alice performing a measurement P^A on a register system A.

2. Alice and Bob make a public announcement, announcing some aspect of their measurement outcomes.

3. Alice and Bob perform post-selection based on these announcements.

4. Alice implements a key map. The key map is a function that maps Alice's raw data and the announcements to a key symbol, chosen from $\{0, 1, ..., N-1\}$ where N is the number of key symbols.

5. Alice performs one-way error correction, leaking some information to Bob, and Bob forms his key.

6. Alice performs privacy amplification, typically by applying a random universal hash function and

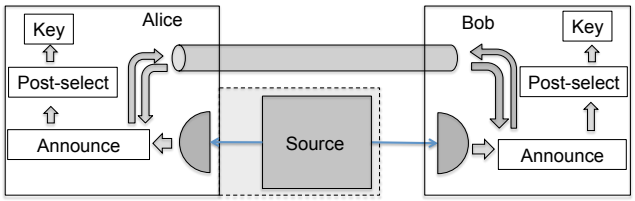

FIG. 2: A prototypical QKD protocol that we use to illustrate our framework. The source sends systems A and B to Alice and Bob, respectively. The dashed line around the source indicates that it may or may not be inside of Alice's laboratory, respectively treating the PM and EB scenarios. Alice and Bob respectively measure POVMs P^A and P^B. Then they publicly announce some information related to their measurement outcomes. These announcements are used to perform post-selection, where some announcement outcomes are discarded. Alice then implements a key map that maps her information (raw data + announcements) to a key variable. Finally, Alice leaks some information (about the result of the key map) to Bob for error correction purposes, and then Bob forms his key.

then communicating the choice of hash function to Bob.

2. *Mathematical model for prototypical protocol*

The quantum state shared by Alice and Bob (prior to their measurements) can be written as ρ_{AB}. To be as pessimistic as possible, we assume that Eve possesses a purification of ρ_{AB}, which we denote as system E. In the following discussion of the protocol, we will note that Eve obtains access to additional systems due to public announcements made by Alice and Bob. (In total, by the end of the protocol, Eve will have access to $E\widetilde{A}\widetilde{B}$ where $\widetilde{A}$ and $\widetilde{B}$ are respectively the registers that store Alice's and Bob's public announcements.)

Consider the experimental constraints on the state ρ_{AB}. These constraints have the form:

$$\mathrm{Tr}((P_j^A \otimes P_k^B)\rho_{AB}) = p_{jk}\,. \tag{35}$$

For prepare-and-measure (PM) protocols, we add additional constraints, as follows. We employ the source-replacement scheme [30, 31], which treats system A as a register that stores the information about which state Alice prepared. This corresponds to Alice preparing the bipartite state

$$|\psi\rangle_{AA'} = \sum_i \sqrt{p_i}|i\rangle_A|\phi_i\rangle_{A'} \tag{36}$$

where $\{|\phi_i\rangle\}$ are the signal states and $\{p_i\}$ are their associated probabilities. Eve's attack maps system A' to system B, producing the state ρ_{AB}. On the other hand, system A is inaccessible to Eve, and hence

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \mathrm{Tr}_{A'}(|\psi\rangle\langle\psi|_{AA'}) \tag{37}$$

is fixed, independent of Eve's attack. To fix ρ_A we add in constraints of the form

$$\mathrm{Tr}((\Theta_j \otimes \mathbb{1}_B)\rho_{AB}) = \theta_j \tag{38}$$

where $\{\Theta_j\}$ is a set of tomographically complete observables on A. Hence, (35) and (38) together represent the constraints that Alice and Bob have on their state.

Step 2 of the above protocol involves Alice and Bob each making an announcement based on their measurement results. In this case, it is helpful to group together the POVM elements into sets associated with particular announcements. Let us write Alice's POVM as $P^A = \{P_j^A\} = \{P^A_{(a,\alpha_a)}\}$ and Bob's POVM as $P^B = \{P_k^B\} = \{P^B_{(b,\beta_b)}\}$. Here the first index denotes the announcement, and the second index denotes a particular element associated with that announcement. Bob's announcement is given by a quantum channel with Kraus operators

$$K_b^B = \sum_{\beta_b} \sqrt{P^B_{(b,\beta_b)}} \otimes |b\rangle_{\widetilde{B}} \otimes |\beta_b\rangle_{\overline{B}}\,. \tag{39}$$

We remark that K_b^B expands the Hilbert space by introducing the registers $\widetilde{B}$ and $\overline{B}$, which is why the operators on these subsystems appear as kets in (39). Similarly, Alice's announcement is given by a quantum channel with Kraus operators

$$K_a^A = \sum_{\alpha_a} \sqrt{P^A_{(a,\alpha_a)}} \otimes |a\rangle_{\widetilde{A}} \otimes |\alpha_a\rangle_{\overline{A}}\,. \tag{40}$$

Here, $\widetilde{A}$ and $\widetilde{B}$ are registers that store Alice's and Bob's announcements, respectively. Also, $\overline{A}$ and $\overline{B}$ are registers that store Alice's and Bob's measurement outcomes, for a given announcement. So, the state after making these announcements becomes

$$\rho^{(2)}_{A\widetilde{A}\overline{A}B\widetilde{B}\overline{B}} = \mathcal{A}(\rho_{AB}) \tag{41}$$

$$= \sum_{a,b} (K_a^A \otimes K_b^B)\rho_{AB}(K_a^A \otimes K_b^B)^\dagger\,, \tag{42}$$

where $\mathcal{A}$ is a CPTP map. We remark that the form of (41) is such that $\widetilde{A}$ and $\widetilde{B}$ are classical registers, meaning that the purifying system has a copy of the registers. This is the way one models a public announcement.

Step 3 is post-selection. Here, Alice and Bob select some announcements to keep and some to discard. Let $\mathbf{A}$ be the set of all announcements that are kept. Then define the projector

$$\Pi = \sum_{(a,b)\in\mathbf{A}} |a\rangle\langle a|_{\widetilde{A}} \otimes |b\rangle\langle b|_{\widetilde{B}}\,, \tag{43}$$

with identity acting on the other systems. The post-selection is modeled by projecting with this projector to obtain the state

$$\rho^{(3)}_{A\widetilde{A}\overline{A}B\widetilde{B}\overline{B}} = \frac{\Pi\rho^{(2)}_{A\widetilde{A}\overline{A}B\widetilde{B}\overline{B}}\Pi}{p_{\mathrm{pass}}}\,. \tag{44}$$

where $p_{\text{pass}} = \text{Tr}(\tilde{\rho}_{A\widetilde{A}\overline{A}B\widetilde{B}\overline{B}}\Pi)$.

In step 4, Alice's chooses a key map. A key map is a function g whose arguments include the outcome of Alice's measurements (a, α_a) and Bob's announcement b. The function outputs a value in $\{0, 1, ..., N-1\}$ where N is the number of key symbols. Hence, we write the key map as the function $g(a, \alpha_a, b)$. We define an isometry V that stores the key information in a register system R, as follows

$$V = \sum_{a,\alpha_a,b} |g(a,\alpha_a,b)\rangle_R \otimes |a\rangle\langle a|_{\widetilde{A}} \otimes |\alpha_a\rangle\langle\alpha_a|_{\overline{A}} \otimes |b\rangle\langle b|_{\widetilde{B}} \,. \tag{45}$$

We first act with this isometry on the state $\rho^{(3)}$ in (44),

$$\rho^{(4)}_{RA\widetilde{A}\overline{A}B\widetilde{B}\overline{B}} = V\rho^{(3)}_{A\widetilde{A}\overline{A}B\widetilde{B}\overline{B}}V^\dagger \,, \tag{46}$$

which stores the key information in the standard basis $\{|j\rangle_R\}$ on R. Then we decohere R in this basis, which turns R into a classical register denoted Z^R, giving the final state

$$\rho^{(5)}_{Z^RA\widetilde{A}\overline{A}B\widetilde{B}\overline{B}} = \mathcal{Z}\left(\rho^{(4)}_{RA\widetilde{A}\overline{A}B\widetilde{B}\overline{B}}\right) , \tag{47}$$

where $\mathcal{Z}$ is a pinching quantum channel, whose action is given by $\mathcal{Z}(\sigma) = \sum_j (|j\rangle\langle j|_R \otimes \mathbb{1})\sigma(|j\rangle\langle j|_R \otimes \mathbb{1})$.

3. Key rate

Finally, Alice performs error correction, which gives $\text{leak}^{\text{EC}}_{\text{obs}}$ number of bits about the key map results to Eve, followed by privacy amplification. This gives the following formula for the key rate [20]:

$$K = p_{\text{pass}}\left[H(Z^R|E\widetilde{A}\widetilde{B})_{\rho^{(5)}} - \text{leak}^{\text{EC}}_{\text{obs}}\right] . \tag{48}$$

Here, $H(A|B)_\rho = H(\rho_{AB}) - H(\rho_B)$ denotes the conditional von Neumann entropy with $H(\sigma) = -\text{Tr}(\sigma\log\sigma)$. As noted above, E is be a purifying system of ρ_{AB}.

More precisely, the expression in (48) must be minimized over all density operators ρ_{AB} that satisfy the constraints in (35) and (38). Hence we write:

$$K = \min_{\rho_{AB}\in\mathbf{S}} \left(p_{\text{pass}} H(Z^R|E\widetilde{A}\widetilde{B})_{\rho^{(5)}}\right) - p_{\text{pass}}\text{leak}^{\text{EC}}_{\text{obs}} \,, \tag{49}$$

where $\mathbf{S}$ has the general form in (4), with the constraints given by (35) and (38).

Having now expressed the key rate in explicit form, we can now relate (49) back to the discussion in Sec. II. This involves simplifying the notation. Namely, for the constraints in (35) and (38), we rewrite them as $\text{Tr}(\Gamma_i\rho) = \gamma_i$, as in (3). Note that we drop the subsystem labels on the state $\rho = \rho_{AB}$. Finally, we write the optimization problem in (49) as

$$\alpha = \min_{\rho\in\mathbf{S}} f(\rho) \,, \tag{50}$$

where

$$f(\rho) = p_{\text{pass}} \cdot H(Z^R|E\widetilde{A}\widetilde{B})_{\rho^{(5)}} \tag{51}$$

$$= p_{\text{pass}} \cdot D\left(\rho^{(4)}_{RA\widetilde{A}\overline{A}B\widetilde{B}\overline{B}} || \rho^{(5)}_{Z^RA\widetilde{A}\overline{A}B\widetilde{B}\overline{B}}\right) \tag{52}$$

$$= D\left(\mathcal{G}(\rho_{AB}) || \mathcal{Z}(\mathcal{G}(\rho_{AB}))\right) . \tag{53}$$

Note that (52) removes the dependence on Eve's system E and is derived using Theorem 1 from [32]. Equation (53) is derived from the previous line using the property $D(c\sigma||c\tau) = cD(\sigma||\tau)$ for any constant $c > 0$. Furthermore we define $\mathcal{G}$ such that its action on an operator σ is given by

$$\mathcal{G}(\sigma) = V\,\Pi\,\mathcal{A}(\sigma)\,\Pi\,V^\dagger \,. \tag{54}$$

Note that (53) has the same form as Eq. (5). In summary, to apply our numerical approach to a given protocol, one formulates $\mathcal{G}$ via (54) and the constraints via (35) and (38). With these objects defined, one then applies our numerical method outlined in Sec. III.

V. EXAMPLES

In this section we consider three practically important scenarios that show the power of our approach. In particular, we consider (1) the BB84 protocol with detector efficiency mismatch, (2) the Trojan horse attack on the BB84 protocol, and (3) the BB84 protocol with phase-coherent signal states. Each of these three scenarios involves a BB84-style protocol but with some "imperfection" accounted for (detector inefficiency, existence of a side channel, and lack of phase randomization). For each scenario, our approach yields significantly higher key rates than those previously obtained in the literature. We remark that the robustness of our numerical approach could allow us to investigate all three imperfections in a single protocol, although for simplicity we consider them separately.

For illustration purposes, for the following examples we assume that error correction is performed at the Shannon limit. This means that the error correction term in (49) can be written as a conditional entropy,

$$\text{leak}^{\text{EC}}_{\text{obs}} = H(Z^R|Z^{\overline{B}}\widetilde{A}\widetilde{B})_{\rho^{(5)}} \,, \tag{55}$$

with the state $\rho^{(5)}$ defined in (47). Here, $Z^{\overline{B}}$ can be viewed as the classical register that one would obtain from measuring in the standard basis on $\overline{B}$.

For each example, further details about the constraints used in our calculations can be found in Appendix E.

A. Efficiency mismatch

Consider a polarization-encoded BB84 protocol, where Bob actively choses his detection basis setting. In this

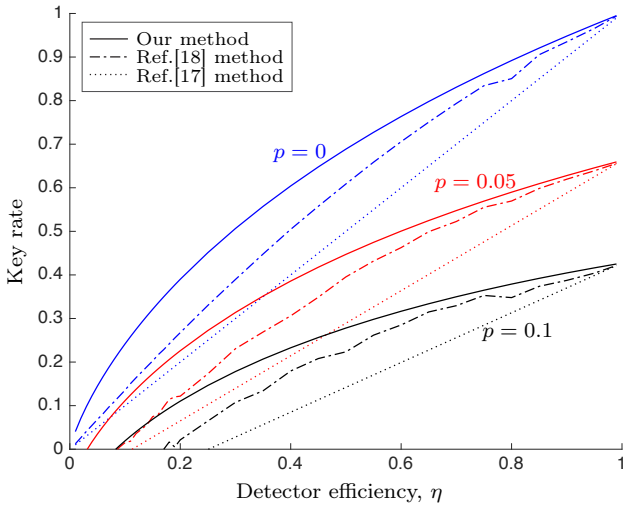

FIG. 3: Key rate for the BB84 protocol with detector efficiency mismatch. Curves are shown for three values of depolarizing probability p (0, 0.05, 0.1). The x-axis is the efficiency of the least efficient detector, with the other detector's efficiency being set to one.

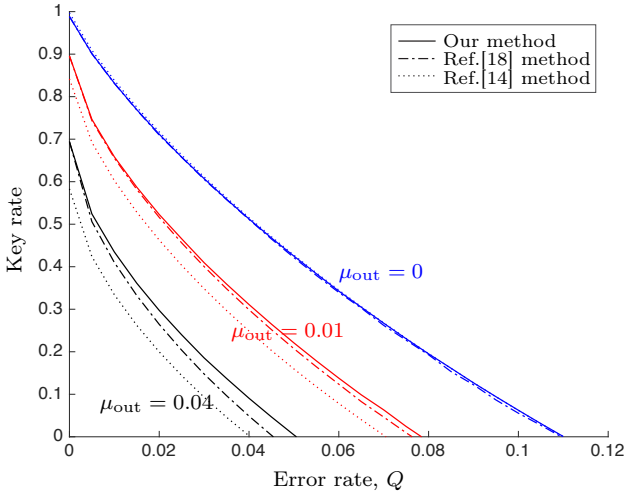

FIG. 4: Key rate vs error rate for the single-photon BB84 protocol under a Trojan-horse attack. The key rate is plotted for different values of μ_{out}. Our numerical method improves on our previous approach in Ref. [18], which in turn gives higher key rates than the analytical method of Ref. [14].

case, Bob's measurement involves two detectors, D_1 and D_2, that are associated with the two polarization states for a given basis.

In practice, it is likely that D_1 and D_2 have different efficiencies, commonly referred to as efficiency mismatch. This mismatch can be further enhanced (by Eve) by manipulating the spatial mode of the incoming light [33]. When efficiency mismatch is large enough, successful hacking strategies on QKD systems have been demonstrated [34]. Furthermore, even with a small amount of efficiency mismatch, the security analysis of QKD becomes difficult to perform.

This motivates the application of our numerical approach the case of detector efficiency mismatch. For simplicity, we consider single-photon signal states, and we assume that no multiple photons arrive at Bob's measurement apparatus. (Our numerical approach can handle multi-photon signals and multi-photon detection events; however, we leave a detailed discussion of the general case for future work.) The single-photon case was previously treated analytically by Fung et al. [17], and hence it provides an opportunity to compare our numerics to the literature.

To further simplify the analysis, we assume one of Bob's detectors is perfect while the other detector has an efficiency η. This allows us to plot the key rate as a function of η, as shown in Fig. 3, for various depolarizing noise levels p. The plot shows that our new numerical method outperforms our previous numerical method based on the dual problem [18], which in turn outperforms the analytical method from Ref. [17].

B. Trojan-horse attack

Consider the phase-encoded BB84 protocol. Here, Alice's light source produces a pulse that passes through an interferometer, one arm of which applies a variable phase θ chosen from the set $\{0, \pi/2, \pi, 3\pi/2\}$ to encode the information. Bob decodes this phase information with an interferometer in his lab.

There is a simple hacking attack on this protocol that exploits a side channel in Alice's encoder (i.e., a channel by which Eve can obtain additional information, beyond the direct channel from Alice to Bob). The attack involves Eve sending a bright pulse of light into Alice's lab [12]. Some fraction of this pulse reaches Alice's phase encoder and is encoded with the same information that Alice is attempting to send to Bob. A portion of this light is reflected back to Eve, who can then decode some of the phase information, potentially compromising the protocol's security. This is called the Trojan-horse attack (THA) [13], since it involves a "malicious gift" from Eve.

For simplicity, let us restrict our attention here to the case where Alice's light source outputs only a single photon per signal. Following the approach of Lucamarini et al. [14], we describe Eve's input pulse and back-reflected pulse as coherent states denoted by $|\sqrt{\mu_{\text{in}}}\rangle$ and $|e^{i\theta}\sqrt{\mu_{\text{out}}}\rangle$, respectively. Here, θ stores Alice's phase encoding setting, and the input and output intensities typically satisfy $\mu_{\text{out}} \ll \mu_{\text{in}}$. The signal states emerging

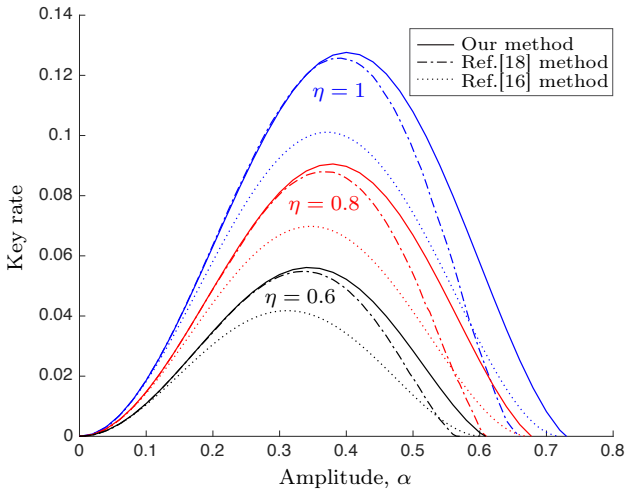

FIG. 5: Key rate versus signal-state amplitude α for three values of transmission probability - $\eta = 1, 0.8, 0.6$ - for the Huttner et al. [15] protocol.

from Alice's source are:

$$|\phi_{x+}\rangle = |x_+\rangle_S \otimes |+\sqrt{\mu_{\text{out}}}\rangle_{S'} \quad (56)$$
$$|\phi_{x-}\rangle = |x_-\rangle_S \otimes |-\sqrt{\mu_{\text{out}}}\rangle_{S'} \quad (57)$$
$$|\phi_{y+}\rangle = |y_+\rangle_S \otimes |+i\sqrt{\mu_{\text{out}}}\rangle_{S'} \quad (58)$$
$$|\phi_{y-}\rangle = |y_-\rangle_S \otimes |-i\sqrt{\mu_{\text{out}}}\rangle_{S'}\,. \quad (59)$$

Here, $|x_\pm\rangle := (1/\sqrt{2})(|z_+\rangle \pm |z_-\rangle)$ and analogously for $|y_\pm\rangle$, where $|z_+\rangle = |1\rangle_L|0\rangle_S$, $|z_-\rangle = |0\rangle_L|1\rangle_S$, and $|n\rangle_L$ ($|n\rangle_S$) is the n-photon state of the long (short) arm of the interferometer.

Figure 4 shows the results of our numerics for the THA. Here we plot key rate versus error rate (assuming the x and y error rates are identical, for simplicity) for various values of μ_{out}. The plot shows that our new numerical method gives higher key rates than both the analytical method from Ref. [14] as well as the numerical method from Ref. [18].

C. BB84 protocol with phase-coherent signal states

Here we consider a protocol proposed by Huttner et al. [15] and analyzed by Lo and Preskill [16]. This is a phase-encoded BB84 protocol, but using coherent states instead of single-photon states. This is quite practical, since one can use attenuated laser pulses from mode-locked lasers to generate these signal states. In addition the protocol is practical because the experimenter does not need to do phase randomization. So it is worth investigating the key rate of this protocol.

The signal states prepared by Alice are [16]

$$|\phi_{z+}\rangle = |+\alpha\rangle_S \otimes |\alpha\rangle_{S'} \quad (60)$$
$$|\phi_{z-}\rangle = |-\alpha\rangle_S \otimes |\alpha\rangle_{S'} \quad (61)$$
$$|\phi_{x+}\rangle = |+i\alpha\rangle_S \otimes |\alpha\rangle_{S'} \quad (62)$$
$$|\phi_{x-}\rangle = |-i\alpha\rangle_S \otimes |\alpha\rangle_{S'} \quad (63)$$

where α is the amplitude of the coherent state, S is the signal mode and S' is the reference mode. When Bob receives the signal, he performs a polarization measurement (in one of two complementary bases), discarding no-click events, and assigning a random bit value to double-click events.

Lo and Preskill [16] gave an analytical lower bound on the key rate for this protocol, as a function of transmission probability η and amplitude α. Their theoretical curves are shown as dotted lines in Fig. 5, for several values of η. In the same plot, we show the result of our numerical optimization as solid lines, with the key rates obtained from the method in Ref. [18] shown as dashed-dotted lines. Interestingly our numerics give higher key rates than the previous literature over the entire parameter range. This is an important result due to the practicality of this protocol.

VI. CONCLUSIONS

In conclusion, we presented a new numerical approach for calculating key rates for QKD. For concreteness, we name our approach the "reliable primal method" or the "reliable primal problem". As discussed in Sec. III, Step 1 of our method is simply the primal optimization problem. Step 2 of our method converts the output of Step 1 (a nearly optimal eavesdropping attack) into a reliable lower bound on the key rate. We presented an efficient method for Step 1 in Sec. III E. Our main contribution is a method for Step 2, which is presented in Theorems 1, 3, and 4.

Reliability is the most important issue with numerical key rate calculations, since key rates must come with a security guarantee. In this work, we highlighted the various issues associated with numerical key rate calculations, such as constraint violation and inexact variable storage by computers. Furthermore, we showed how to address these issues. Our most robust result, Theorem 4, allows one to lower bound the key rate despite numerical imprecision.

We discussed that our method is arbitrarily tight in Sec III D. This allowed us to make significant improvements over previous literature key rates for three interesting examples in Sec. V. Furthermore, the tightness of our approach implies that the solid curves that we plotted in Figs. 3, 4, and 5 are essentially unbeatable, i.e., they cannot be improved upon. Eliminating looseness from key rate calculations is a major advance for the field of QKD research.

Future applications of our work include investigating device imperfections, side channels, multi-photon detection events, decoy-state protocols with partial phase randomization, measurement-device independent protocols, differential-phase shift protocols, and coherent one-way protocols. Perhaps more importantly, our approach can allow researchers to explore and evaluate novel protocol ideas that have yet to be discovered.

As noted in the Introduction, our group released a user-friendly software for key rate calculations based on the dual problem from Ref. [18]. Interestingly, our reliable primal method presented here improves on the approach of Ref. [18] in terms of both speed and tightness. Therefore, we plan to improve our publicly-available software in the future by incorporating the reliable primal method. We believe this software has the potential to be used throughout the QKD community, both in industry and academia.

Finally, we hope to extend our approach to finite-key analysis [9, 35] in the near future.

VII. ACKNOWLEDGEMENTS

We acknowledge support from Industry Canada, Sandia National Laboratories, Office of Naval Research (ONR), NSERC Discovery Grant, and Ontario Research Fund (ORF).

[1] Campagna, M. *et al.* *Quantum Safe Cryptography and Security* (European Telecommunications Standards Institute, 2015).

[2] Wyner, A. D. The WireTap Channel. *Bell System Technical Journal* **54**, 1355–1387 (1975).

[3] Scarani, V. *et al.* The security of practical quantum key distribution. *Reviews of Modern Physics* **81**, 1301–1350 (2009).

[4] Lo, H.-K., Curty, M. & Tamaki, K. Secure quantum key distribution. *Nature Photonics* **8**, 595–604 (2014).

[5] Xin, H. Space science. Chinese Academy takes space under its wing. *Science (New York, N.Y.)* **332**, 904 (2011).

[6] Peev, M. *et al.* The SECOQC quantum key distribution network in Vienna. *New Journal of Physics* **11**, 075001 (2009).

[7] Sasaki, M. *et al.* Field test of quantum key distribution in the Tokyo QKD Network. *Optics Express* **19**, 10387–10409 (2011).

[8] Wang, S. *et al.* Field and long-term demonstration of a wide area quantum key distribution network. *Optics Express* **22**, 329–342 (2014).

[9] Renner, R. *Security of Quantum Key Distribution*. Ph.D. thesis, ETH Zurich (2005).

[10] Renner, R., Gisin, N. & Kraus, B. Information-theoretic security proof for quantum-key-distribution protocols. *Physical Review A* **72**, 012332 (2005).

[11] Watanabe, S., Matsumoto, R. & Uyematsu, T. Tomography increases key rates of quantum-key-distribution protocols. *Physical Review A* **78**, 042316 (2008).

[12] Vakhitov, A., Makarov, V. & Hjelme, D. R. Large pulse attack as a method of conventional optical eavesdropping in quantum cryptography. *Journal of Modern Optics* **48**, 2023–2038 (2001).

[13] Gisin, N., Fasel, S., Kraus, B., Zbinden, H. & Ribordy, G. Trojan-horse attacks on quantum-key-distribution systems. *Physical Review A* **73**, 022320 (2006).

[14] Lucamarini, M. *et al.* Practical security bounds against the trojan-horse attack in quantum key distribution. *Physical Review X* **5**, 1–19 (2015).

[15] Huttner, B., Imoto, N., Gisin, N. & Mor, T. Quantum cryptography with coherent states. *Physical Review A* **51**, 1863–1869 (1995). 9502020.

[16] Lo, H. & Preskill, J. Security of quantum key distribution using weak coherent states with nonrandom phases. *Quantum Information and Computation* **7**, 431–458 (2007).

[17] Fung, C. C.-H. F., Tamaki, K., Qi, B., Lo, H.-K. H. & Ma, X. Security proof of quantum key distribution with detection efficiency mismatch. *Quantum Inf. ˜Comput.* **9**, 0131–0165 (2009). 0802.3788.

[18] Coles, P. J., Metodiev, E. M. & Lütkenhaus, N. Numerical approach for unstructured quantum key distribution. *Nature Communications* **7**, 11712 (2016). 1510.01294.

[19] Boyd, S. & Vandenberghe, L. *Convex Optimization* (Cambridge University Press, 2004).

[20] Devetak, I. & Winter, A. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A* **461**, 207–235 (2005).

[21] Al-Mohy, A. H. & Higham, N. J. Improved Inverse Scaling and Squaring Algorithms for the Matrix Logarithm IMPROVED INVERSE SCALING AND SQUARING ALGORITHMS FOR THE MATRIX LOGARITHM *. *SIAM J. Sci. Comput.* **34**, 153–169 (2012).

[22] Frank, M. & Wolfe, P. An algorithm for quadratic programming. *Naval Research Logistics Quarterly* **3**, 95–110 (1956).

[23] Bennett, C. H. & Brassard, G. Quantum cryptography: public key distribution and coin tossing. In *International Conference on Computers, Systems & Signal Processing, Bangalore, India*, 175–179 (1984).

[24] Bennett, C. H. Quantum cryptography using any two nonorthogonal states. *Physical Review Letters* **68**, 3121–3124 (1992).

[25] Scarani, V., A. Acin, G. R., Gisin, N., Acín, A. & Ribordy, G. Quantum Cryptography Protocols Robust against Photon Number Splitting Attacks for Weak Laser Pulse Implementations. *Physical review letters* **92**, 057901 (2004). 0211131.

[26] Bruss, D. Optimal eavesdropping in quantum cryptography with six states. *Physical Review Letters* **81**, 3018–3021 (1998).

[27] Lo, H.-K., Ma, X. & Chen, K. Decoy State Quantum Key Distribution. *Physical Review Letters* **94**, 230504 (2005).

[28] Lo, H.-K., Curty, M. & Qi, B. Measurement-Device-Independent Quantum Key Distribution. *Physical Review Letters* **108**, 130503 (2012).

[29] Stacey, W., Annabestani, R., Ma, X. & Lütkenhaus, N. Security of quantum key distribution using a simplified trusted relay. *Physical Review A* **91**, 012338 (2015). 1408.4426.

[30] Bennett, C., Brassard, G. & Mermin, N. Quantum cryptography without Bell's theorem. *Physical Review Letters* **68**, 557–559 (1992).

[31] Ferenczi, A. & Lütkenhaus, N. Symmetries in quantum key distribution and the connection between optimal attacks and optimal cloning. *Physical Review A* **85**, 052310 (2012).

[32] Coles, P. J. Unification of different views of decoherence and discord. *Physical Review A* **85**, 042103 (2012).

[33] Sajeed, S. *et al.* Security loophole in free-space quantum key distribution due to spatial-mode detector-efficiency mismatch. *Physical Review A - Atomic, Molecular, and Optical Physics* **91**, 1–6 (2015). 1502.02785.

[34] Lydersen, L. *et al.* Hacking commercial quantum cryptography systems by tailored bright illumination. *Nature Photonics* **4**, 5 (2010).

[35] Scarani, V. & Renner, R. Quantum cryptography with finite resources: Unconditional security bound for discrete-variable protocols with one-way postprocessing. *Physical Review Letters* **100**, 200501 (2008).

[36] Nielsen, M. A. & Chuang, I. *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[37] Lo, H.-K., Chau, H. F. & Ardehali, M. Efficient quantum key distribution scheme and a proof of its unconditional security. *J. Cryptology* **18**, 133–165 (2004).

[38] Gittsovich, O. *et al.* Squashing model for detectors and applications to quantum-key-distribution protocols. *Physical Review A* **89**, 1–29 (2014). 1310.5059.

Appendix A: Proof of Theorem 1

1. Standard form for semidefinite programs

To prove Theorem 1, we will make use of the so-called standard form for semidefinite programs (e.g., see page 265 of [19]), given as follows:

Primal Problem	Dual Problem	
maximize $\langle A, X\rangle$	minimize $\vec{\gamma}\cdot\vec{y}$	(A2)
subject to:	subject to:	(A3)
$\langle B_1, X\rangle = \gamma_1$	$\vec{y}\cdot\vec{B} \geqslant A$	(A4)
$\vdots$	$\vec{y}\in\mathbb{R}^n$	(A5)
$\langle B_m, X\rangle = \gamma_m$		(A6)
$X \geqslant 0$		(A7)

Here, the inner product is of the Hilbert-Schmidt form, $\langle A, B\rangle := \mathrm{Tr}(A^\dagger B)$, and the vector notation $\vec{a}\cdot\vec{b}$ is shorthand for $\sum_{j=1}^m a_j b_j$.

2. Inequality in (8)

In what follows, we first prove the inequality in (8), and in the next subsection we establish the equality condition.

For some matrix σ, let $\vec{\sigma} := \mathrm{vec}(\sigma)$ denote the vectorization obtained by stacking the columns of σ. For two matrices σ and τ, note that the inner product between their vectorizations can be written as

$$\vec{\sigma}\cdot\vec{\tau} = \mathrm{Tr}(\sigma^T\tau)\,, \tag{A8}$$

where ^{T} is the transpose taken in the basis used to define the vectorization.

Now consider the gradient matrix $\nabla f(\rho)$ defined in (7) and its vectorization $\vec{\nabla} f(\rho)$. Note that the quantity

$$g(\rho,\sigma) := (\vec{\sigma}-\vec{\rho})\cdot\vec{\nabla} f(\rho) \tag{A9}$$

$$= \mathrm{Tr}\left[(\sigma-\rho)^T\nabla f(\rho)\right] \tag{A10}$$

quantifies how much the function f changes whenever one moves from point ρ to point σ. [Note that Eq. (A10) rewrote $g(\rho,\sigma)$ in matrix notation, where ^{T} is the transpose in the same basis that is used to represent the gradient matrix.] More precisely, $g(\rho,\sigma)$ quantifies how much the linearization L_ρ of f changes, where L_ρ is the linearization about point ρ. Specifically, the linearization is given by

$$L_\rho(\sigma) = f(\rho) + g(\rho,\sigma)\,. \tag{A11}$$

Since f is a convex, differentiable function over a convex set $\mathbf{S}$, the linearization L_ρ always lies below the curve f (e.g. page 69 of [19]). Hence, for any two points $\rho, \sigma \in \mathbf{S}$ we have

$$f(\sigma) - f(\rho) \geqslant g(\rho, \sigma)\,. \tag{A12}$$

Now let $\rho^* \in \mathbf{S}$ minimize f over $\mathbf{S}$. Then

$$\begin{aligned} f(\rho^*) &\geqslant f(\rho) + \mathrm{Tr}((\rho^* - \rho)^T \nabla f(\rho)) && \text{(A13)}\\ &\geqslant f(\rho) + \min_{\sigma \in \mathbf{S}} \left[\mathrm{Tr}((\sigma - \rho)^T \nabla f(\rho))\right] && \text{(A14)}\\ &= f(\rho) - \mathrm{Tr}(\rho^T \nabla f(\rho)) + \min_{\sigma \in \mathbf{S}} \mathrm{Tr}(\sigma^T \nabla f(\rho))\,, && \text{(A15)} \end{aligned}$$

where (A14) exploits the fact that $\rho^* \in \mathbf{S}$. Hence finding a lower bound on α reduces to the minimization problem

$$\min_{\sigma \in \mathbf{S}} \mathrm{Tr}(\sigma^T \nabla f(\rho)). \tag{A16}$$

This is a linear semidefinite program (SDP) and we may apply duality theory to obtain the dual SDP. In particular, our problem is essential the standard SDP form given in Sec. A 1, which gives the following dual problem

$$\max_{\vec{y}\ \in \mathbf{S}^*(\rho)} \vec{\gamma} \cdot \vec{y}\,, \tag{A17}$$

where

$$\mathbf{S} = \{\rho \in \mathbf{H}_+ \mid \mathrm{Tr}(\Gamma_i \rho) = \gamma_i, \forall i\}\,, \tag{A18}$$

$$\mathbf{S}^*(\sigma) = \left\{ \vec{y} \in \mathbb{R}^n \mid \sum_i y_i \Gamma_i^T \leqslant \nabla f(\sigma) \right\}. \tag{A19}$$

Weak duality implies that

$$\min_{\sigma \in \mathbf{S}} \mathrm{Tr}(\sigma^T \nabla f(\rho)) \geqslant \max_{\vec{y}\ \in \mathbf{S}^*(\rho)} \vec{\gamma} \cdot \vec{y}\,. \tag{A20}$$

Inserting (A20) into (A15) gives the desired lower bound in (8).

3. Equality in (8)

With the inequality in Theorem 1 proven, we now turn to establishing equality in (8) if $f(\rho) = f(\rho^*)$.

First, consider the inequality in (A20). Slater's condition provides a sufficient criteria for strong duality to hold (e.g. page 265 of [19]).

In general, Slater's condition only applies to the primal problem rather than the dual. However, in the case of a semidefinite program, we remark that the dual (A16) and primal (A17) problem are in direct correspondence. Hence, we can apply Slater's conditon with the primal and dual problem interchanged.

To satisfy the condition it is adequate to show that $\mathbf{S} \neq \emptyset$ and there exists $\vec{y} \in \mathbb{R}^n$ such that $\sum_i y_i \Gamma_i^T < \nabla f(\rho)$. Since the set of constraints $\{\Gamma_i\}$ correspond to a valid density matrix, it immediately follows that $\mathbf{S} \neq \emptyset$. Since density matrices are constrained to have trace one, without loss of generality we may take $\Gamma_1 = \mathbb{1}$ and $\gamma_1 = 1$. Thus, if $\lambda_{\min}$ is the smallest eigenvalue of $\nabla f(\rho)$, it follows that $(\lambda_{\min} - 1)\Gamma_1^T < \nabla f(\rho)$. So $\vec{y} = (\lambda_{\min} - 1, 0, ..., 0)^T$ satisfies $\sum_i y_i \Gamma_i^T < \nabla f(\rho)$. With Slater's condition satisfied, strong duality holds and

$$\min_{\sigma \in \mathbf{S}} \mathrm{Tr}(\sigma^T \nabla f(\rho)) = \max_{\vec{y}\ \in \mathbf{S}^*(\rho)} \vec{\gamma} \cdot \vec{y}\,. \tag{A21}$$

We now show that if $f(\rho) = f(\rho^*)$ then equality in (8) holds. Suppose that $f(\rho) = f(\rho^*)$, then (A15) yields

$$\min_{\sigma \in \mathbf{S}} \mathrm{Tr}((\sigma - \rho)^T \nabla f(\rho)) \leqslant 0\,. \tag{A22}$$

Next, we state a lemma that provides a bound in the opposite direction of (A22).

Lemma 5: For a point ρ that minimizes f over $\mathbf{S}$,

$$\min_{\sigma \in \mathbf{S}} \mathrm{Tr}((\sigma - \rho)^T \nabla f(\rho)) \geqslant 0\,. \tag{A23}$$

Proof. The Karush-Kuhn-Tucker (KKT) conditions (e.g. page 243 of [19]) provide necessary conditions for the optimality of ρ. For our problem they say that if ρ is optimal, then there exists a pair $(\vec{\lambda}, Z) \in \mathbb{R}^n \times \mathbf{H}^{d\times d}$ such that

$$\nabla f(\rho) + \sum_i \lambda_i {\Gamma_i}^T - Z = 0\,, \tag{A24}$$

$$\mathrm{Tr}(Z^T\rho) = 0\,, \tag{A25}$$

$$Z \geqslant 0\,, \tag{A26}$$

where we have differentiated our constraints to get the latter two terms in (A24). Let $\sigma \in \mathbf{S}$ then

$$\mathrm{Tr}((\sigma-\rho)^T\nabla f(\rho)) = \mathrm{Tr}((\sigma-\rho)^T(-\sum_i \lambda_i {\Gamma_i}^T + Z)) \tag{A27}$$

$$= \sum_i \lambda_i \mathrm{Tr}(\Gamma_i\rho) - \sum_i \lambda_i \mathrm{Tr}(\Gamma_i\sigma) + \mathrm{Tr}(Z^T\sigma) - \mathrm{Tr}(Z^T\rho) \tag{A28}$$

$$= \mathrm{Tr}(Z^T\sigma)\,, \tag{A29}$$

where we have used the definition of $\mathbf{S}$ and (A25) in the last equality. Since σ and Z^T are both positive semidefinite it follows that the trace of their product is nonnegative. Hence,

$$\mathrm{Tr}((\sigma-\rho)^T\nabla f(\rho)) \geqslant 0\,, \tag{A30}$$

and since σ is arbitrary, the desired result follows. □

Combining (A22) and (A23) we must have

$$\min_{\sigma\in\mathbf{S}} \mathrm{Tr}((\sigma-\rho)^T\nabla f(\rho)) = 0\,. \tag{A31}$$

Consequently,

$$f(\rho^*) = f(\rho) + \min_{\sigma\in\mathbf{S}} \mathrm{Tr}((\sigma-\rho)^T\nabla f(\rho)) \tag{A32}$$

$$= f(\rho) - \mathrm{Tr}(\rho^T\nabla f(\rho)) + \max_{\vec{y}\in\mathbf{S}^*(\rho)} \vec{\gamma}\cdot\vec{y} \tag{A33}$$

$$= \beta(\rho)\,. \tag{A34}$$

Appendix B: Existence of the gradient for $\rho > 0$

1. Some useful lemmas

Here we show that if $\rho > 0$, i.e., if ρ is full rank, then the gradient given by Eq. (11) is well defined. To justify this statement, we first state several useful lemmas.

Lemma 6: Let $\mathcal{E}$ be a quantum channel, and let Π denote the projector onto the support of $\mathcal{E}(\mathbb{1})$. Then for any operator X,

$$\mathcal{E}(X) = \Pi\mathcal{E}(X)\Pi \tag{B1}$$

Proof. Let P be a positive semidefinite operator such that $0 \leqslant P \leqslant \mathbb{1}$, then $\mathcal{E}(P) \leqslant \mathcal{E}(\mathbb{1})$. This implies that $\mathrm{supp}(\mathcal{E}(P)) \subseteq \mathrm{supp}(\mathcal{E}(\mathbb{1}))$. In turn this implies that

$$\mathcal{E}(P) = \Pi\mathcal{E}(P)\Pi. \tag{B2}$$

Note that one can multiply the above equation by any positive number p and it still holds. Hence defining $Q = pP$ we obtain

$$\mathcal{E}(Q) = \Pi\mathcal{E}(Q)\Pi. \tag{B3}$$

for any $Q \geqslant 0$. Since the positive operators form a basis for the operator space, any operator X can be written as a linear combination of positive operators. Hence, taking linear combinations of equations of the form of (B3), we arrive at the desired result (B1). □

Lemma 7: Let $\mathcal{E}$ be a quantum channel, and let Π denote the projector onto the support of $\mathcal{E}(\mathbb{1})$. Then for any operator Y,

$$\mathcal{E}^\dagger(Y) = \mathcal{E}^\dagger(\Pi Y \Pi) \tag{B4}$$

where $\mathcal{E}^\dagger$ is the adjoint of $\mathcal{E}$.

Proof. Let $\langle A, B\rangle = \mathrm{Tr}(A^\dagger B)$ be the Hilbert-Schmidt inner product. Consider some operator X. Then

$$\langle \mathcal{E}^\dagger(Y), X\rangle = \langle Y, \mathcal{E}(X)\rangle = \langle Y, \Pi\mathcal{E}(X)\Pi\rangle = \langle \Pi Y\Pi, \mathcal{E}(X)\rangle = \langle \mathcal{E}^\dagger(\Pi Y\Pi), X\rangle, \tag{B5}$$

where we invoked (B1). Since X is arbitrary, we have $\mathcal{E}^\dagger(Y) = \mathcal{E}^\dagger(\Pi Y\Pi)$. □

Lemma 8: Let $\mathcal{E}$ be a quantum channel, then for any full-rank density matrix $\rho > 0$,

$$\mathrm{supp}(\mathcal{E}(\rho)) = \mathrm{supp}(\mathcal{E}(\mathbb{1}))\,. \tag{B6}$$

Proof. The fact that $\rho > 0$ implies that there exists an $\epsilon > 0$ such that

$$\epsilon\mathbb{1} \leqslant \rho \leqslant \mathbb{1}\,, \tag{B7}$$

where the second inequality follows from ρ being a density matrix. Since $\mathcal{E}$ is completely positive,

$$\epsilon\mathcal{E}(\mathbb{1}) \leqslant \mathcal{E}(\rho) \leqslant \mathcal{E}(\mathbb{1})\,, \tag{B8}$$

and it is clear that (B6) follows from this equation. □

Lemma 9: Let $\mathcal{E}$ be a pinching quantum channel, i.e., one whose action is given by $\mathcal{E}(\rho) = \sum_j \Pi_j\rho\Pi_j$ where the Π_j are orthogonal projectors such that $\sum_j \Pi_j = \mathbb{1}$. Then for any $\rho \geqslant 0$,

$$\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\mathcal{E}(\rho))\,. \tag{B9}$$

Proof. One way to see this is to consider the von Neumann entropy

$$\begin{aligned} H(\mathcal{E}(\rho)) &= -\mathrm{Tr}\left[\mathcal{E}(\rho)\log\mathcal{E}(\rho)\right] &\text{(B10)}\\ &= -\mathrm{Tr}\left[\rho\mathcal{E}(\log\mathcal{E}(\rho))\right] &\text{(B11)}\\ &= -\mathrm{Tr}\left[\rho\log\mathcal{E}(\rho)\right]\,, &\text{(B12)}\end{aligned}$$

where (B12) follows because $\log\mathcal{E}(\rho)$ is already pinched, and so pinching it again with $\mathcal{E}$ has no effect. The quantity in (B12) would be ill-defined if and only if ρ is not contained inside $\mathrm{supp}(\mathcal{E}(\rho))$. In contrast, the von Neumann entropy in (B10) is a well-behaved function that does not blow up under any circumstances. Since the two quantities are equal, this must mean that ρ is contained inside $\mathrm{supp}(\mathcal{E}(\rho))$. □

2. Proof of Lemma 2

As noted in (11), the gradient has the form

$$[\nabla f(\rho)]^T = \mathcal{G}^\dagger(\log\mathcal{G}(\rho)) - \mathcal{G}^\dagger(\log\mathcal{Z}(\mathcal{G}(\rho)))\,. \tag{B13}$$

The quantities $\mathcal{G}(\rho)$ and $\mathcal{Z}(\mathcal{G}(\rho))$ may be singular and hence the conventional matrix logarithm is ill-defined. We remedy the situation by redefining the matrix logarithm with an extension that holds even when $\mathcal{G}(\rho)$ or $\mathcal{Z}(\mathcal{G}(\rho))$ are singular.

We are only interested in the case where the argument of the matrix logarithm σ is hermitian and hence, we may diagonalize it as $\sigma = P^{-1}DP$ where P is an orthogonal matrix and $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. A conventional way of evaluating the matrix logarithm for hermitian $\sigma > 0$ is

$$\log\sigma = P^{-1}(\log D)P = P^{-1}\mathrm{diag}(\ln\lambda_1, \ln\lambda_2, \ldots, \ln\lambda_n)P \tag{B14}$$

We extend this definition of the matrix logarithm to all $\sigma \geqslant 0$ by redefining it as

$$\log\sigma := P^{-1}\mathrm{diag}(\tilde{\ln}\lambda_1, \tilde{\ln}\lambda_2, \ldots, \tilde{\ln}\lambda_n)P \tag{B15}$$

where $\tilde{\ln} : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ is defined by

$$\tilde{\ln}(x) := \begin{cases} \ln(x)\,, & x > 0 \\ 0\,, & x = 0 \end{cases} \tag{B16}$$

With our new definition of the matrix logarithm, the gradient is well defined for all $\rho > 0$ since any zero eigenvalues of $\mathcal{G}(\rho)$ and $\mathcal{Z}(\mathcal{G}(\rho))$ are mapped to 0 by $\tilde{\ln}$. In fact, mapping the zero eigenvalues to any real number by $\tilde{\ln}$ will prove a valid extension.

Note that the new definition of the gradient is only different from the standard definition when $\mathcal{G}(\rho)$ or $\mathcal{Z}(\mathcal{G}(\rho))$ has a zero eigenvalue, so we consider those two cases.

Consider the first term in (B13). The zero eigenvalues in $\mathcal{G}(\rho)$ correspond to eigenvectors outside the support of $\mathcal{G}$ by definition. Let $\lambda_i = 0$ be once such eigenvalue with corresponding eigenvector $|\lambda_i\rangle$. Suppose for a moment that $\lambda_i \neq 0$. From our definition of the matrix logarithm, it is apparent that $\log \mathcal{G}(\rho)$ has an eigenvalue $\ln \lambda_i$ with corresponding eigenvector $|\lambda_i\rangle$. By Lemma 7, it follows that this subspace is anhilated when $\mathcal{G}^\dagger$ is applied to $\log \mathcal{G}(\rho)$. Hence the value assigned to $\ln \lambda_i$ is irrelevant so long as it is a number. Now consider $\lambda_i = 0$. Our previous discussion implies that for all $\lambda_i \neq 0$ the associated subspace is destroyed. We only run into trouble when $\lambda_i = 0$.To repair the situation, we define $\tilde{\ln}(0) = 0$ since the subspace plays no role in the value of the gradient. This is precisely our extension of the first term.

Likewise, combining Lemma 7 and 9 yields an identical line of reasoning that justifies the extension of the second term.

3. Continuity

Here we state that the objective function $f(\rho)$ is continuous, which will be a useful lemma needed for our proof of Theorem 3. First we state the following lemma for the trace distance (or trace norm).

Lemma 10: Let ρ and σ be (normalized) density matrices. Let $\mathcal{E}$ be a completely positive trace non-increasing (CPTNI) map. Then

$$\|\rho - \sigma\|_1 \geqslant \|\mathcal{E}(\rho) - \mathcal{E}(\sigma)\|_1\,, \tag{B17}$$

where $\|A\|_1 = \mathrm{Tr}\sqrt{A^\dagger A}$ is the trace norm.

The proof is a straightforward extension of the proof for CPTP maps found in [36].

Next we state a lemma, known as Fannes' inequality, that entropy is continuous, whose proof can also be found in [36].

Lemma 11: Let $\rho \geqslant 0$ and $\sigma \geqslant 0$ be $d \times d$ matrices, such that $\|\rho - \sigma\|_1 \leqslant \epsilon \leqslant 1/e$. Then

$$|H(\rho) - H(\sigma)| \leqslant \epsilon \log(d/\epsilon)\,. \tag{B18}$$

Note that the right-hand side of (B18) goes to zero as $\epsilon \to 0$. Finally we state the continuity of our objective function.

Lemma 12: Let ρ and σ be (normalized) density matrices such that $\|\rho - \sigma\|_1 \leqslant \epsilon \leqslant 1/e$. Let $f(\rho)$ be defined as in (5),

$$f(\rho) = D(\mathcal{G}(\rho)||\mathcal{Z}(\mathcal{G}(\rho))\,, \tag{B19}$$

where $\mathcal{G}$ is a completely positive map and $\mathcal{Z}$ is a pinching quantum channel. Suppose $\mathcal{G}(\rho)$ and $\mathcal{G}(\sigma)$ are $d \times d$. Then

$$|f(\rho) - f(\sigma)| \leqslant 2\epsilon \log(d/\epsilon)\,. \tag{B20}$$

Proof. It is straightforward to show that

$$f(\rho) = H\left[\mathcal{Z}(\mathcal{G}(\rho))\right] - H\left[\mathcal{G}(\rho)\right]\,. \tag{B21}$$

Next note that Lemma 10 implies that $\|\mathcal{G}(\rho) - \mathcal{G}(\sigma)\|_1 \leqslant \epsilon$ and $\|\mathcal{Z}(\mathcal{G}(\rho)) - \mathcal{Z}(\mathcal{G}(\sigma))\|_1 \leqslant \epsilon$. So from (B21) we have

$$\begin{aligned} |f(\rho) - f(\sigma)| &= |H\left[\mathcal{Z}(\mathcal{G}(\rho))\right] - H\left[\mathcal{G}(\rho)\right] - H\left[\mathcal{Z}(\mathcal{G}(\sigma))\right] + H\left[\mathcal{G}(\sigma)\right]| && \text{(B22)} \\ &\leqslant |H\left[\mathcal{Z}(\mathcal{G}(\rho))\right] - H\left[\mathcal{Z}(\mathcal{G}(\sigma))\right]\,| + |H\left[\mathcal{G}(\rho)\right] - H\left[\mathcal{G}(\sigma)\right]\,| && \text{(B23)} \\ &\leqslant \epsilon \log(d/\epsilon) + \epsilon \log(d/\epsilon) && \text{(B24)} \\ &\leqslant 2\epsilon \log(d/\epsilon)\,, && \text{(B25)} \end{aligned}$$

where (B24) invoked Lemma 11. □

4. Proof of Theorem 3

We first define a perturbed optimization problem. Let $\epsilon \in \mathbb{R}$ such that $0 < \epsilon < 1/(de)$, then consider

$$\alpha(\epsilon) := \min_{\rho \in \mathbf{S}(\epsilon)} f(\rho) \tag{B26}$$

where

$$\mathbf{S}(\epsilon) := \{\rho \geqslant \epsilon \mathbb{1} \mid \mathrm{Tr}(\Gamma_i \rho) = \widehat{\gamma}_i, \forall i\}\,, \tag{B27}$$

$$\widehat{\gamma}_i := \gamma_i + \epsilon \mathrm{Tr}(\Gamma_i)\,. \tag{B28}$$

There is a natural bijection $\mathcal{M} : \mathbf{S} \to \mathbf{S}(\epsilon)$ where

$$\mathcal{M}(\rho) = \rho + \epsilon \mathbb{1}\,. \tag{B29}$$

Note that this bijection is operationally identical to defining $\rho(\epsilon) := \rho + \epsilon \mathbb{1}$, as we did in Theorem 3.

Let $\rho(\epsilon) \in \mathbf{S}(\epsilon)$. If $\rho(\epsilon)^*$ minimizes f over $\mathbf{S}(\epsilon)$ then we can apply the proof of Theorem 1 to obtain

$$f(\rho(\epsilon)^*) \geqslant f(\rho(\epsilon)) - \mathrm{Tr}\left[\rho(\epsilon)^T \nabla f(\rho(\epsilon))\right] + \min_{\sigma \in \mathbf{S}(\epsilon)} \mathrm{Tr}\left[\sigma^T \nabla f(\rho(\epsilon))\right]\,. \tag{B30}$$

The optimization problem in (B30) is

$$\min_{\sigma \in \mathbf{H}} \mathrm{Tr}\left[\sigma^T \nabla f(\rho(\epsilon))\right] \tag{B31}$$

$$\text{s.t.} \mathrm{Tr}(\Gamma_i \sigma) = \widehat{\gamma}_i \tag{B32}$$

$$\sigma \geqslant \epsilon \mathbb{1} \tag{B33}$$

where $\mathbf{H}$ denotes the set of hermitian $d \times d$ matrices. Since $\mathcal{M}$ is a bijection, we can define $\tilde{\sigma} = \sigma - \epsilon \mathbb{1}$ and rewrite the optimization problem as

$$\min_{\tilde{\sigma} \in \mathbf{H}} \mathrm{Tr}\left[(\tilde{\sigma} + \epsilon \mathbb{1})^T \nabla f(\rho(\epsilon))\right] \tag{B34}$$

$$\text{s.t.} \mathrm{Tr}\left[\Gamma_i(\tilde{\sigma} + \epsilon \mathbb{1})\right] = \widehat{\gamma}_i \tag{B35}$$

$$\tilde{\sigma} + \epsilon \mathbb{1} \geqslant \epsilon \mathbb{1} \tag{B36}$$

Then simplify to get

$$\min_{\tilde{\sigma} \in \mathbf{H}} \mathrm{Tr}\left[\tilde{\sigma}^T \nabla f(\rho(\epsilon))\right] + \epsilon \mathrm{Tr}\left[\nabla f(\rho(\epsilon))\right] \tag{B37}$$

$$\text{s.t.} \mathrm{Tr}(\Gamma_i \tilde{\sigma}) = \gamma_i \tag{B38}$$

$$\tilde{\sigma} \geqslant 0 \tag{B39}$$

The first term in (B37) is identical to the optimization problem in Theorem 1. The latter term is a constant, so it follows that

$$f(\rho(\epsilon)^*) \geqslant \beta(\rho(\epsilon)) + \epsilon \mathrm{Tr}(\nabla f(\rho(\epsilon))) \tag{B40}$$

where equality holds if $f(\rho(\epsilon)) = f(\rho(\epsilon)^*)$.

Next we establish a relationship between $f(\rho^*)$ and $f(\rho(\epsilon)^*)$. We can apply Lemma 12 to obtain

$$|f(\mathcal{M}_\epsilon(\rho)) - f(\rho)| \leqslant -2d\epsilon \log_2 \epsilon = \delta_\epsilon \tag{B41}$$

for any $\rho \in \mathbf{S}$. If ρ^* minimizes f over $\mathbf{S}$, it follows that

$$f(\rho^*) \geqslant f(\mathcal{M}(\rho^*)) - \delta_\epsilon\,. \tag{B42}$$

Thus, if $\rho(\epsilon)^*$ minimizes f over $\mathbf{S}(\epsilon)$ then

$$f(\rho^*) \geqslant f(\rho(\epsilon)^*) - \delta_\epsilon\,. \tag{B43}$$

Hence we have

$$\alpha \geqslant f(\rho(\epsilon)^*) - \delta_\epsilon\,. \tag{B44}$$

For any $\rho(\epsilon) \in \mathbf{S}(\epsilon)$, we can apply (B40) to obtain

$$\alpha \geqslant \beta(\rho(\epsilon)) + \epsilon \mathrm{Tr}(\nabla f(\rho(\epsilon))) - \delta_\epsilon\,. \tag{B45}$$

Appendix C: Computational implementation of the lower-bounding method

1. Handling imprecise representations

As noted in the main text (Sec. III C), it is impossible to provide exact floating-point representations of $\{\Gamma_i\}$ and $\{\gamma_i\}$. We address this impossibility by defining approximate representations, which we denote by $\{\tilde{\Gamma}_i\}$ and $\{\tilde{\gamma}_i\}$ respectively. We relate the approximate and exact representations by

$$\tilde{\Gamma}_i = \Gamma_i + \delta\Gamma_i \quad \text{and} \quad \tilde{\gamma}_i = \gamma_i + \delta\gamma_i \,, \tag{C1}$$

where

$$\|\delta\Gamma_i\|_{\text{HS}} < \epsilon_1 \quad \text{and} \quad |\delta\gamma_i| < \epsilon_2 \,, \tag{C2}$$

for all i. Here, the Hilbert-Schmidt norm is defined by $\|A\|_{\text{HS}} := \sqrt{\text{Tr}(A^\dagger A)}$. Determining the constants ϵ_1 and ϵ_2 is rather technical and depends heavily on how the approximate observables and expectation values are computed. However, if the mantissa of the underlying representation is increased in length (or arbitrary precision arithmetic is used) and appropriate numerical algorithms are applied, the approximate representations will become arbitrarily accuracte. Hence, ϵ_1 and ϵ_2 may be made as small as needed.

We can define the quantity

$$\epsilon_{\text{rep}} = \epsilon_1 + \epsilon_2 \tag{C3}$$

that measures our overall variable uncertainty. We now prove a lemma that motivates our treatment of the imprecision.
Lemma 13: Let $\rho \in \mathbf{S}$ and let the definitions in (C1), (C2), and (C3) hold. Then

$$|\text{Tr}(\tilde{\Gamma}_i\rho) - \tilde{\gamma}_i| < \epsilon_{\text{rep}} \,. \tag{C4}$$

Proof. We can apply the triangle inequality, the Cauchy-Schwarz inequality and the fact that ρ is a density matrix to obtain

$$\begin{aligned}
|\text{Tr}(\tilde{\Gamma}_i\rho) - \tilde{\gamma}_i| &= |\text{Tr}(\delta\Gamma_i\rho) - \delta\gamma_i| && \text{(C5)}\\
&\leqslant |\text{Tr}(\delta\Gamma_i\rho)| + |\delta\gamma_i| && \text{(C6)}\\
&\leqslant \|\delta\Gamma_i\|_{\text{HS}}\|\rho\|_{\text{HS}} + |\delta\gamma_i| && \text{(C7)}\\
&\leqslant \|\delta\Gamma_i\|_{\text{HS}} + |\delta\gamma_i| && \text{(C8)}\\
&< \epsilon_1 + \epsilon_2 \,. && \text{(C9)}
\end{aligned}$$

□

2. Handling imprecise solvers

Up to this point in our error analysis, we have neglected the fact that no numerical solver is exact. In reality, the matrix $\tilde{\rho}$ returned by the solver may not be positive semidefinite or satisfy the approximate constraints defined by $\{\tilde{\Gamma}_i\}$ and $\{\tilde{\gamma}_i\}$. In what follows we present a method for addressing this situation.

Let λ_{min} denote the smallest eigenvalue of $\tilde{\rho}$, then

$$\tilde{\rho}' := \begin{cases} \tilde{\rho} - \lambda_{\text{min}}\mathbb{1} \,, & \lambda_{\text{min}} < 0 \\ \tilde{\rho} \,, & \text{otherwise} \end{cases} \tag{C10}$$

is positive semidefinite. Let ϵ_{sol} be a positive real number such that

$$|\text{Tr}(\tilde{\Gamma}_i\tilde{\rho}') - \tilde{\gamma}_i| < \epsilon_{\text{sol}} \,. \tag{C11}$$

The quantity ϵ_{sol} describes how close $\tilde{\rho}'$ is to satisfying the approximate constraints provided that it is chosen to be as small as possible. The closer $\tilde{\rho}'$ is to conforming to the constraints, the smaller ϵ_{sol} will be. So

We now have two quantities: ϵ_{rep} describes the representation precision and ϵ_{sol} describes the solver precision. We define the quantity

$$\epsilon' = \max(\epsilon_{\text{rep}}, \epsilon_{\text{sol}}) \,. \tag{C12}$$

Recall the relaxed set of approximate density matrices defined in (18),

$$\mathbf{S}_{\epsilon'} = \left\{ \rho \in \mathbf{H}_+ \middle| |\mathrm{Tr}(\tilde{\Gamma}_i \rho) - \tilde{\gamma}_i| < \epsilon', \forall i \right\} . \tag{C13}$$

From Lemma 13, it follows that $\mathbf{S} \subseteq \mathbf{S}_{\epsilon'}$ and our previous discussion implies that $\tilde{\rho}' \in \mathbf{S}_{\epsilon'}$. Furthermore,

$$\lim_{\epsilon' \to 0+} \mathbf{S}_{\epsilon'} = \mathbf{S} , \tag{C14}$$

so it is apparent that $\mathbf{S}_{\epsilon'}$ is a natural generalization of $\mathbf{S}$.

3. Proof of Theorem 4

Recall that

$$\alpha = \min_{\rho \in \mathbf{S}} f(\rho) , \tag{C15}$$

$$\alpha_{\epsilon'} = \min_{\rho \in \mathbf{S}_{\epsilon'}} f(\rho) . \tag{C16}$$

Note that since $\mathbf{S} \subseteq \mathbf{S}_{\epsilon'}$ it follows that $\alpha_{\epsilon'} \leqslant \alpha$. This means that any lower bound on $\alpha_{\epsilon'}$ is also a lower bound on α.

Next we consider a small perturbation as in Theorem 3. By applying an argument identical to that in Theorem 3 and invoking the relation $\alpha_{\epsilon'} \leqslant \alpha$ we find that for any $\rho \in \mathbf{S}_{\epsilon'}$ and $0 < \epsilon < 1/(de)$

$$\alpha \geqslant f(\rho) - \mathrm{Tr}(\rho^T \nabla f(\rho)) + \epsilon \mathrm{Tr}(\nabla f(\rho)) - \delta_\epsilon + \min_{\sigma \in \mathbf{S}_{\epsilon'}} \mathrm{Tr}(\sigma^T \nabla f(\rho)) . \tag{C17}$$

We now focus on transforming the minimization in (C17) into a maximization via duality theory. First we rewrite the minimization problem as

$$\min_{\sigma \in \mathbf{H}} \mathrm{Tr}(\sigma^T \nabla f(\rho)) \tag{C18}$$

$$\text{s.t.} \mathrm{Tr}(\tilde{\Gamma}_i \sigma) \leqslant \tilde{\gamma}_i + \epsilon' \tag{C19}$$

$$\mathrm{Tr}(-\tilde{\Gamma}_i \sigma) \leqslant -\tilde{\gamma}_i + \epsilon' \tag{C20}$$

$$\sigma \geqslant 0 \tag{C21}$$

We introduce the slack variables $\vec{a}$ and $\vec{b}$ so that the problem becomes

$$\min_{\sigma \in \mathbf{H}} \mathrm{Tr}(\sigma^T \nabla f(\rho)) \tag{C22}$$

$$\text{s.t.} \mathrm{Tr}(\tilde{\Gamma}_i \sigma) + a_i = \tilde{\gamma}_i + \epsilon' \tag{C23}$$

$$\mathrm{Tr}(-\tilde{\Gamma}_i \sigma) + b_i = -\tilde{\gamma}_i + \epsilon' \tag{C24}$$

$$\sigma \geqslant 0 \tag{C25}$$

$$\vec{a}, \vec{b} \geqslant 0 \tag{C26}$$

Next we recast the problem so that there is again one positive semidefinite variable. Define the block-diagonal matrices

$$\tilde{\sigma} = \mathrm{diag}(\sigma, \vec{a}^T, \vec{b}^T) \tag{C27}$$

$$\nabla \tilde{f}(\rho) = \mathrm{diag}(\nabla f(\rho), \vec{0}) \tag{C28}$$

$$\tilde{\Gamma}_i^+ = \mathrm{diag}(\tilde{\Gamma}_i, \delta_{i1}, \delta_{i2}, ..., \delta_{in}, \vec{0}^T) \tag{C29}$$

$$\tilde{\Gamma}_i^- = \mathrm{diag}(-\tilde{\Gamma}_i, \vec{0}^T, \delta_{i1}, \delta_{i2}, ..., \delta_{in}) \tag{C30}$$

where δ_{ij} denotes the Kronecker delta. The optimization problem then becomes

$$\min_{\tilde{\sigma} \in \mathbf{H}} \mathrm{Tr}(\tilde{\sigma}^T \nabla \tilde{f}(\rho)) \tag{C31}$$

$$\text{s.t.} \mathrm{Tr}(\tilde{\Gamma}_i^{\pm} \tilde{\sigma}) = \pm \tilde{\gamma}_i + \epsilon' \tag{C32}$$

$$\tilde{\sigma} \geqslant 0 \tag{C33}$$

This is a semidefinite program of a form identical to the one that appears in the proof of Theorem 1. Duality theory yields the dual problem

$$\max_{\vec{y}\in\mathbf{S}^*_{\epsilon'}(\rho)} (\vec{\tilde{\gamma}}^T + \epsilon', -\vec{\tilde{\gamma}}^T + \epsilon')^T \cdot \vec{y}. \tag{C34}$$

Substituting (C34) into (C17) gives the desired lower bound.

Appendix D: Tightness

1. Some useful lemmas

Here we state some lemmas that will eventually help us prove the tightness of our lower bounding method. We begin with a corollary of Theorem 3, as follows.

Corollary 14: If $\rho(\epsilon)^*$ minimizes f over $\mathbf{S}(\epsilon)$ and ρ^* minimizes f over $\mathbf{S}$ then

$$|f(\rho(\epsilon)^*) - f(\rho^*)| \leqslant \delta_\epsilon \tag{D1}$$

Proof. We first note that (D1) can be rewritten as

$$f(\rho(\epsilon)^*) - \delta_\epsilon \leqslant f(\rho^*) \leqslant f(\rho(\epsilon)^*) + \delta_\epsilon \tag{D2}$$

where the leftmost inequality was proven in (B43) and hence, we only need to show that

$$f(\rho^*) \leqslant f(\rho(\epsilon)^*) + \delta_\epsilon \,. \tag{D3}$$

We may invoke (B41) to obtain

$$f(\sigma) \leqslant f(\mathcal{M}(\sigma)) + \delta_\epsilon \tag{D4}$$

where $\sigma = \mathcal{M}^{-1}(\rho(\epsilon)^*)$. Thus

$$f(\mathcal{M}^{-1}(\rho(\epsilon)^*)) \leqslant f(\rho(\epsilon)^*) + \delta_\epsilon \tag{D5}$$

Since $\rho(\epsilon)$ minimizes f over $\mathbf{S}$ and $\mathcal{M}^{-1} : \mathbf{S}(\epsilon) \to \mathbf{S}$ it follows that

$$f(\rho^*) \leqslant f(\rho(\epsilon)^*) + \delta_\epsilon \,. \tag{D6}$$

□

Next we state a few lemmas that pertain to the trace of the gradient, starting with the following.

Lemma 15: For a $d \times d$ density matrix σ, define $\sigma(\epsilon) = \sigma + \epsilon\mathbb{1}$ with $\epsilon > 0$. Then,

$$\lim_{\epsilon\to 0} \epsilon \mathrm{Tr}\left[\log(\sigma(\epsilon))\right] = 0 \,. \tag{D7}$$

Proof. Since $0 \leqslant \sigma \leqslant \mathbb{1}$, it follows that

$$\epsilon\mathbb{1} \leqslant \sigma(\epsilon) \leqslant (1+\epsilon)\mathbb{1} \,. \tag{D8}$$

Now note that the matrix logarithm is an operator monotone, hence

$$(\log\epsilon)\mathbb{1} \leqslant \log(\sigma(\epsilon)) \leqslant (\log(1+\epsilon))\mathbb{1} \,. \tag{D9}$$

Taking the trace and multiplying by ϵ gives

$$d\epsilon\log\epsilon \leqslant \epsilon \mathrm{Tr}\left[\log(\sigma(\epsilon))\right] \leqslant d\epsilon\log(1+\epsilon) \,. \tag{D10}$$

Now taking the limit $\epsilon \to 0$ we see that both the left and right sides of (D10) go to zero, proving the desired result. □

Now we state a slightly more general lemma, whose proof is similar.

Lemma 16: For a $d \times d$ density matrix σ, define $\sigma(\epsilon) = \sigma + \epsilon \mathbb{1}$ with $\epsilon > 0$. Also, let $\mathcal{E}$ and $\mathcal{F}$ be completely positive (CP) maps that are independent of ϵ. Then,

$$\lim_{\epsilon \to 0} \epsilon \mathrm{Tr}\left[\mathcal{E}(\log(\mathcal{F}(\sigma(\epsilon))))\right] = 0\,. \tag{D11}$$

Proof. Follow the same proof as the previous lemma, starting with (D8). Next note that $\mathcal{F}$ is CP, so it preserves positivity, giving

$$\epsilon \mathcal{F}(\mathbb{1}) \leqslant \mathcal{F}(\sigma(\epsilon)) \leqslant (1+\epsilon)\mathcal{F}(\mathbb{1})\,. \tag{D12}$$

Taking the logarithm and using the fact that $\mathcal{E}$ is CP gives

$$\mathcal{E}(\log(\epsilon \mathcal{F}(\mathbb{1}))) \leqslant \mathcal{E}(\log(\mathcal{F}(\sigma(\epsilon)))) \leqslant \mathcal{E}(\log((1+\epsilon)\mathcal{F}(\mathbb{1})))\,. \tag{D13}$$

Taking the trace and multiplying by epsilon gives

$$L(\epsilon) \leqslant \epsilon \mathrm{Tr}\left[\mathcal{E}(\log(\mathcal{F}(\sigma(\epsilon))))\right] \leqslant R(\epsilon) \tag{D14}$$

with

$$L(\epsilon) := \epsilon \mathrm{Tr}\left[\mathcal{E}(\log(\epsilon \mathcal{F}(\mathbb{1})))\right], \quad R(\epsilon) := \epsilon \mathrm{Tr}\left[\mathcal{E}(\log((1+\epsilon)\mathcal{F}(\mathbb{1})))\right]\,. \tag{D15}$$

Note that for a matrix A and constant c, we have $\log(cA) = (\log c)\mathbb{1} + \log A$. This gives

$$\begin{aligned} L(\epsilon) &= \epsilon \mathrm{Tr}\left[\mathcal{E}((\log \epsilon)\mathbb{1} + \log(\mathcal{F}(\mathbb{1})))\right] && \text{(D16)} \\ &= (\epsilon \log \epsilon)\mathrm{Tr}\left[\mathcal{E}(\mathbb{1})\right] + \epsilon \mathrm{Tr}\left[\mathcal{E}(\log(\mathcal{F}(\mathbb{1})))\right]\,. && \text{(D17)} \end{aligned}$$

Since both $\mathcal{E}$ and $\mathcal{F}$ are independent of ϵ, both the first and second term of (D17) go to zero as $\epsilon \to 0$, and hence

$$\lim_{\epsilon \to 0} L(\epsilon) = 0. \tag{D18}$$

Similarly

$$R(\epsilon) = (\epsilon \log(1+\epsilon))\mathrm{Tr}\left[\mathcal{E}(\mathbb{1})\right] + \epsilon \mathrm{Tr}\left[\mathcal{E}(\log(\mathcal{F}(\mathbb{1})))\right]\,, \quad \text{and} \quad \lim_{\epsilon \to 0} R(\epsilon) = 0. \tag{D19}$$

This proves the desired result. □

Proposition 17: For any $d \times d$ density matrix σ, define $\sigma(\epsilon) = \sigma + \epsilon \mathbb{1}$. The gradient formula in (11) satisfies

$$\lim_{\epsilon \to 0} \epsilon \mathrm{Tr}\left[\nabla f(\sigma(\epsilon))\right] = 0\,. \tag{D20}$$

Proof. From Eq. (11), we have

$$\begin{aligned} \epsilon \mathrm{Tr}\left[\nabla f(\sigma(\epsilon))\right] &= T_1(\epsilon) - T_2(\epsilon), \quad \text{with} && \text{(D21)} \\ T_1(\epsilon) &:= \epsilon \mathrm{Tr}\left[\mathcal{G}^\dagger(\log \mathcal{G}(\sigma(\epsilon)))\right] && \text{(D22)} \\ T_2(\epsilon) &:= \epsilon \mathrm{Tr}\left[\mathcal{G}^\dagger(\log \mathcal{Z}(\mathcal{G}(\sigma(\epsilon))))\right]\,. && \text{(D23)} \end{aligned}$$

Note that both $T_1(\epsilon)$ and $T_2(\epsilon)$ have the same form as the expression appearing in Lemma 16. For $T_1(\epsilon)$, apply Lemma 16 with $\mathcal{E} = \mathcal{G}^\dagger$ and $\mathcal{F} = \mathcal{G}$ to find that

$$\lim_{\epsilon \to 0} T_1(\epsilon) = 0\,. \tag{D24}$$

Likewise for $T_2(\epsilon)$, apply Lemma 16 with $\mathcal{E} = \mathcal{G}^\dagger$ and $\mathcal{F} = \mathcal{Z} \circ \mathcal{G}$ to find that

$$\lim_{\epsilon \to 0} T_2(\epsilon) = 0\,. \tag{D25}$$

Combining (D24) and (D25) gives the desired result. □

2. Tightness of the lowerbound in Theorem 4

In this section we show that the method in Theorem 4 is tight. We want to show that as $\epsilon \to 0$ and $\epsilon' \to 0$, the minimizer $\rho^*_{\epsilon'}$ of f over $\mathbf{S}_{\epsilon'}$ satisfies

$$\beta_{\epsilon\epsilon'}(\rho^*_{\epsilon'}(\epsilon)) - \delta_\epsilon \to f(\rho^*) \tag{D26}$$

where ρ^* minimizes f over $\mathbf{S}$ and the left-hand side of (D26) is the lower bound produced by Theorem 4. We state this as the following theorem.

Theorem 18: Let $\rho^*_{\epsilon'}$ be the minimizer of f over $\mathbf{S}_{\epsilon'}$. Then

$$\lim_{\epsilon,\epsilon'\to 0} \beta_{\epsilon\epsilon'}(\rho^*_{\epsilon'}(\epsilon)) - \delta_\epsilon = f(\rho^*)\,. \tag{D27}$$

Proof. Taking the limit of the lower bound in Theorem 4 as $\epsilon' \to 0$ we recover the bound in Theorem 3. That is

$$\lim_{\epsilon,\epsilon'\to 0} \beta_{\epsilon\epsilon'}(\rho^*_{\epsilon'}(\epsilon)) - \delta_\epsilon = \lim_{\epsilon\to 0} \beta_\epsilon(\rho^*(\epsilon)) - \delta_\epsilon\,. \tag{D28}$$

Now by (B40)

$$\lim_{\epsilon,\epsilon'\to 0} \beta_{\epsilon\epsilon'}(\rho^*_{\epsilon'}(\epsilon)) - \delta_\epsilon = \lim_{\epsilon\to 0} f(\rho^*(\epsilon)) - \epsilon \mathrm{Tr}(\nabla f(\rho^*(\epsilon))) - \delta_\epsilon \tag{D29}$$

$$= \lim_{\epsilon\to 0} \left[f(\rho^*(\epsilon)) - \delta_\epsilon\right] - \lim_{\epsilon\to 0} \left[\epsilon \mathrm{Tr}(\nabla f(\rho^*(\epsilon)))\right]\,. \tag{D30}$$

Note that the second term in (D30) involving the gradient vanishes, from Proposition 17. This gives

$$\lim_{\epsilon,\epsilon'\to 0} \beta_{\epsilon\epsilon'}(\rho^*_{\epsilon'}(\epsilon)) - \delta_\epsilon = \lim_{\epsilon\to 0} \left[f(\rho^*(\epsilon)) - \delta_\epsilon\right]\,. \tag{D31}$$

Finally, Corollary 14 implies that

$$\lim_{\epsilon,\epsilon'\to 0} \beta_{\epsilon\epsilon'}(\rho^*_{\epsilon'}(\epsilon)) - \delta_\epsilon = f(\rho^*)\,. \tag{D32}$$

□

Thus, given suitably small ϵ, ϵ' it follows that the bound produced by Theorem 4 applied to a near-optimal state is arbitrarily tight.

Appendix E: Examples

1. Efficiency Mismatch

For the BB84 protocol with detector efficiency mismatch, we model it as an entanglement-based protocol. We write Alice's POVM as

$$P_1^A = p_z|0\rangle\langle 0|\,, \quad P_2^A = p_z|1\rangle\langle 1|\,, \quad P_3^A = (1-p_z)|+\rangle\langle +|\,, \quad P_4^A = (1-p_z)|-\rangle\langle -| \tag{E1}$$

where $\{|0\rangle, |1\rangle\}$ is the z-basis on a qubit, and $|\pm\rangle = 1/\sqrt{2}(|0\rangle \pm |1\rangle)$. Here p_z denotes the probability for Alice to measure in the z-basis. For our numerics, we chose $p_z \approx 1$, corresponding to using the z-basis most of the time.

We model Bob's system as a qutrit, where the one-photon subspace is modeled as a qubit subspace, and the third dimension is the vacuum. This third dimension is incorporated because of detector inefficiency, which may cause a no-click event. Bob's POVM elements associated with detecting a photon are given by

$$P_1^B = p_z|0\rangle\langle 0| \oplus 0\,, \quad P_2^B = p_z\eta|1\rangle\langle 1| \oplus 0\,, \quad P_3^B = (1-p_z)|+\rangle\langle +| \oplus 0\,, \quad P_4^B = (1-p_z)\eta|-\rangle\langle -| \oplus 0\,, \tag{E2}$$

where the direct sum is used here to embed qubit operators inside a qutrit Hilbert space. Note that P_2^B and P_4^B have a factor of η due to detector inefficiency. (We assume one detector has perfect efficiency, while the other has efficiency η.) The last of Bob's POVM elements corresponds to a no-click event,

$$P_5^B = \mathbb{1} - \sum_{j=1}^{4} P_j^B\,. \tag{E3}$$

To generate the constraints in (35), we simulate the data using a depolarizing channel with depolarizing probability p,

$$\mathcal{E}_{\text{dep}}(\rho) = (1-p)\rho + p\mathbb{1}/2\,. \tag{E4}$$

We consider the bipartite state generated from applying this channel to half of a maximally entangled state $|\Phi\rangle$,

$$\rho_{AB}^{\text{sim}} = (\mathcal{I}\otimes\mathcal{E}_{\text{dep}})(|\Phi\rangle\langle\Phi|)\,, \tag{E5}$$

and we emphasize that this state is only used to simulate experimental data. To obtain the constraints in (35), we compute

$$p_{jk} = \text{Tr}((P_j^A\otimes P_k^B)\rho_{AB}^{\text{sim}})\,. \tag{E6}$$

Now consider Alice's and Bob's announcements. For sifting purposes, Alice announces her basis, and so (40) becomes

$$K_0^A = \sqrt{P_1^A}\otimes|0\rangle_{\widetilde{A}}\otimes|0\rangle_{\overline{A}} + \sqrt{P_2^A}\otimes|0\rangle_{\widetilde{A}}\otimes|1\rangle_{\overline{A}} \tag{E7}$$

$$K_1^A = \sqrt{P_3^A}\otimes|1\rangle_{\widetilde{A}}\otimes|0\rangle_{\overline{A}} + \sqrt{P_4^A}\otimes|1\rangle_{\widetilde{A}}\otimes|1\rangle_{\overline{A}}\,. \tag{E8}$$

Likewise Bob announces his basis and he announces whether he got a click or not. We can model this with three Kraus operators as follows

$$K_0^B = \sqrt{P_1^B}\otimes|0\rangle_{\widetilde{B}}\otimes|0\rangle_{\overline{B}} + \sqrt{P_2^B}\otimes|0\rangle_{\widetilde{B}}\otimes|1\rangle_{\overline{B}} \tag{E9}$$

$$K_1^B = \sqrt{P_3^B}\otimes|1\rangle_{\widetilde{B}}\otimes|0\rangle_{\overline{B}} + \sqrt{P_4^B}\otimes|1\rangle_{\widetilde{B}}\otimes|1\rangle_{\overline{B}} \tag{E10}$$

$$K_2^B = \sqrt{P_5^B}\otimes|2\rangle_{\widetilde{B}}\otimes|0\rangle_{\overline{B}}\,. \tag{E11}$$

Next we consider the post-selection. Events where Bob does not receive a click, or where Alice and Bob use different bases, are discarded. Hence, (43) becomes

$$\Pi = |0\rangle\langle 0|_{\widetilde{A}}\otimes|0\rangle\langle 0|_{\widetilde{B}} + |1\rangle\langle 1|_{\widetilde{A}}\otimes|1\rangle\langle 1|_{\widetilde{B}}\,. \tag{E12}$$

Finally, consider the isometry V associated with the key map defined in (45). We can define the key map such that Alice stores 0 (1) in her key when she obtains outcome P_1^A or P_3^A (P_2^A or P_4^A). This gives

$$V = |0\rangle_R\otimes|0\rangle\langle 0|_{\overline{A}} + |1\rangle_R\otimes|1\rangle\langle 1|_{\overline{A}}\,, \tag{E13}$$

with identity acting on all other subsystems. The above expressions allow one to define $\mathcal{G}$ in (54), and hence define the optimization problem.

2. Trojan-horse attack

We model the BB84 protocol under a Trojan-horse attack as a prepare-and-measure protocol with sifting. As discussed in Sec. IV, we treat this by constructing the source-replacement state,

$$|\psi\rangle_{AA'} = \sqrt{\frac{p_x}{2}}|0\rangle|\phi_{x+}\rangle + \sqrt{\frac{p_x}{2}}|1\rangle|\phi_{x-}\rangle + \sqrt{\frac{1-p_x}{2}}|2\rangle|\phi_{y+}\rangle + \sqrt{\frac{1-p_x}{2}}|3\rangle|\phi_{y-}\rangle \tag{E14}$$

where $\{|\phi_{x\pm}\rangle, |\phi_{y\pm}\rangle\}$ are the signal states specified in (56)-(59). For high-efficiency sifting [37], we bias the probability distribution so that the x-basis is used most of the time, i.e., $p_x \approx 1$. Within this framework, Alice prepares her signal states by acting with a POVM on register system A, with POVM elements

$$P_1^A = |0\rangle\langle 0|\,,\quad P_2^A = |1\rangle\langle 1|\,,\quad P_3^A = |2\rangle\langle 2|\,,\quad P_4^A = |3\rangle\langle 3|\,. \tag{E15}$$

Bob measures in either the x- or y-basis via the following POVM

$$P_1^B = p_x|x_+\rangle\langle x_+|\,,\quad P_2^B = p_x|x_-\rangle\langle x_-|\,,\quad P_3^B = (1-p_x)|y_+\rangle\langle y_+|\,,\quad P_4^B = (1-p_x)|y_-\rangle\langle y_-|\,, \tag{E16}$$

where for simplicity we set the p_x appearing in Bob's measurement to be the same value as that used for Alice's signal states.

Next we consider data simulation for the purpose of formulating the constraints in (35). We model Eve's attack as a depolarizing channel (E4) with depolarizing probability p. (Note that $p = 2Q$, where Q is the error rate plotted in Fig. 4.) Applying this channel to the state $|\psi\rangle_{AA'}$ gives

$$\rho_{AB}^{\text{sim}} = (\mathcal{I} \otimes \mathcal{E}_{\text{dep}})(|\psi\rangle\langle\psi|_{AA'}) \,. \tag{E17}$$

To obtain the constraints in (35), we compute

$$p_{jk} = \text{Tr}((P_j^A \otimes P_k^B)\rho_{AB}^{\text{sim}}) \,. \tag{E18}$$

Since Alice's density operator is fixed, we add the additional constraints specified by (38).

Now we consider the announcements made by Alice and Bob. Alice announces her choice of basis, so (40) becomes

$$K_0^A = \sqrt{P_1^A} \otimes |0\rangle_{\widetilde{A}} \otimes |0\rangle_{\overline{A}} + \sqrt{P_2^A} \otimes |0\rangle_{\widetilde{A}} \otimes |1\rangle_{\overline{A}} \tag{E19}$$

$$K_1^A = \sqrt{P_3^A} \otimes |1\rangle_{\widetilde{A}} \otimes |0\rangle_{\overline{A}} + \sqrt{P_4^A} \otimes |1\rangle_{\widetilde{A}} \otimes |1\rangle_{\overline{A}} \,. \tag{E20}$$

(We remark that, in this case, introducing the additional register system $\overline{A}$ is redundant since the key information can be read off directly from system A, but we do it here for completeness.)

Similarly, Bob announces his choice of basis, so (39) becomes

$$K_0^B = \sqrt{P_1^B} \otimes |0\rangle_{\widetilde{B}} \otimes |0\rangle_{\overline{B}} + \sqrt{P_2^B} \otimes |0\rangle_{\widetilde{B}} \otimes |1\rangle_{\overline{B}} \tag{E21}$$

$$K_1^B = \sqrt{P_3^B} \otimes |1\rangle_{\widetilde{B}} \otimes |0\rangle_{\overline{B}} + \sqrt{P_4^B} \otimes |1\rangle_{\widetilde{B}} \otimes |1\rangle_{\overline{B}} \,. \tag{E22}$$

For the post-selection, Alice and Bob discard events where they measure in different bases. So (43) becomes

$$\Pi = |0\rangle\langle 0|_{\widetilde{A}} \otimes |0\rangle\langle 0|_{\widetilde{B}} + |1\rangle\langle 1|_{\widetilde{A}} \otimes |1\rangle\langle 1|_{\widetilde{B}} \,. \tag{E23}$$

Finally, consider the isometry V associated with the key map defined in (45). We can define the key map such that Alice stores 0 (1) in her key when she obtains outcome P_1^A or P_3^A (P_2^A or P_4^A). This gives

$$V = |0\rangle_R \otimes |0\rangle\langle 0|_{\overline{A}} + |1\rangle_R \otimes |1\rangle\langle 1|_{\overline{A}} \,, \tag{E24}$$

with identity acting on all other subsystems. The above expressions allow one to define $\mathcal{G}$ in (54), and hence define the optimization problem.

3. BB84 protocol with phase-coherent signal states

We model the BB84 protocol with phase-coherent signal states as a prepare-and-measure protocol with sifting, similar to how we modeled the Trojan-horse attack above. We apply the source-replacement scheme as described in Sec. IV, with the state $|\psi\rangle_{AA'}$ from (36) given by

$$|\psi\rangle_{AA'} = \sqrt{\frac{p_z}{2}}|0\rangle|\phi_{z+}\rangle + \sqrt{\frac{p_z}{2}}|1\rangle|\phi_{z-}\rangle + \sqrt{\frac{1-p_z}{2}}|2\rangle|\phi_{x+}\rangle + \sqrt{\frac{1-p_z}{2}}|3\rangle|\phi_{x-}\rangle \tag{E25}$$

where $\{|\phi_{z\pm}\rangle, |\phi_{x\pm}\rangle\}$ are specified in (60)-(63). Here, p_z denotes the probability of Alice preparing a state in the z-basis, and it is biased to be close to one.

Alice's POVM acts on the register system A with the standard basis elements, namely

$$P_1^A = |0\rangle\langle 0| \,, \quad P_2^A = |1\rangle\langle 1| \,, \quad P_3^A = |2\rangle\langle 2| \,, \quad P_4^A = |3\rangle\langle 3| \,. \tag{E26}$$

By applying a squashing model [38], we model Bob's system as a qutrit, where the one-photon subspace is modeled as a qubit subspace, and the third dimension is the vacuum. This third dimension is incorporated because of channel loss, which may cause a no-click event. Bob's POVM elements are then

$$P_1^B = p_z|0\rangle\langle 0| \oplus 0 \,, \quad P_2^B = p_z|1\rangle\langle 1| \oplus 0 \,, \quad P_3^B = (1-p_z)|+\rangle\langle +| \oplus 0 \,, \quad P_4^B = (1-p_z)|-\rangle\langle -| \oplus 0 \,, \quad P_5^B = \mathbb{1} - \sum_{i=1}^{4} P_i^B \,. \tag{E27}$$

Next we consider data simulation for the purpose of formulating the constraints in (35). We model the channel between Alice and Bob as a lossy channel $\mathcal{E}_{\text{loss}}(\rho)$ with transmission probability η. Note that the action of this channel on a coherent state is $|\alpha\rangle \to |\sqrt{\eta}\alpha\rangle$. We can apply the channel to the state $|\psi\rangle_{AA'}$ to obtain the bipartite state

$$\rho_{AB}^{\text{sim}} = (\mathcal{I} \otimes \mathcal{E}_{\text{loss}})(|\psi\rangle\langle\psi|_{AA'}) \,. \tag{E28}$$

The remainder of the model is identical to that of Sec. E 2 from (E18) onwards. This defines the optimization problem.