Loss-tolerant quantum key distribution with mixed signal states

J. Eli Bourassa,[1,*] Ignatius William Primaatmaja,[2] Charles Ci Wen Lim,[2,3,†] and Hoi-Kwong Lo[1,4,5,6]

[1]*Department of Physics, University of Toronto, Toronto, Canada*
[2]*Centre for Quantum Technologies, National University of Singapore, Singapore*
[3]*Department of Electrical and Computer Engineering, National University of Singapore, Singapore*
[4]*Center for Quantum Information and Quantum Control, University of Toronto, Toronto, Canada*
[5]*Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada*
[6]*Department of Physics, University of Hong Kong, Hong Kong*

(Received 27 July 2020; accepted 17 November 2020; published 9 December 2020)

The security of measurement device-independent quantum key distribution (MDI QKD) relies on a thorough characterization of one's optical source output, especially any noise in the state preparation process. Here, we provide an extension of the loss-tolerant protocol [Phys. Rev. A **90**, 052314 (2014)], a leading proof technique for analyzing the security of QKD, to MDI QKD protocols that employ mixed signal states. We first reframe the core of the proof technique, noting its generalization to treat d-dimensional signal encodings. Concentrating on the qubit signal state case, we find that the mixed states can be interpreted as providing Alice and Bob with a virtual shield system they can employ to reduce Eve's knowledge of the secret key. We then introduce a simple semidefinite programming method for optimizing the virtual twisting operations they can perform on the shield system to yield a higher key rate, along with an example calculation of fundamentally achievable key rates in the case of random polarization modulation error.

DOI: 10.1103/PhysRevA.102.062607

I. INTRODUCTION

There has been significant interest in quantum hacking against practical quantum key distribution (QKD) systems [1,2]. In particular, single photon detectors (SPDs) have been identified as the weakest link in the security of practical QKD. To completely bypass all possible attacks on SPDs, the concept of measurement-device-independent (MDI) QKD has been introduced and widely deployed. MDI QKD allows two distant parties, Alice and Bob, to distribute a shared, secret cryptographic key, even in the presence of an eavesdropper, Eve, who has complete control of their quantum channels and measurement devices [3,4]. Typically, Alice and Bob prepare a set of signal states, send them to a central measurement node potentially controlled by Eve, which then makes an announcement based on a measurement it may not have faithfully executed. The cost of information-theoretic security in this setting is that Alice and Bob need to trust and characterize the optical sources they employ to send signals. Thus, it is especially valuable to account for the source features and flaws in a security proof when quantifying the key rates offered by an MDI protocol.

In this work, we answer a seemingly simple question: How does one construct a security proof for an MDI QKD protocol that employs trusted, yet noisy—i.e., mixed—signal states? To clarify, protocols that employ the decoy state method [5,6] call for mixed optical states in the form of phase-randomized weak coherent pulses. However, in those protocols, the *signal* states—i.e., the single photon contributions—used for key generation are still often assumed to be pure. Unfortunately, a realistic source will not be able to initialize signal states with perfect purity. Therefore, our task is to build a consistent framework for optimally determining the security of MDI QKD protocols in the case of mixed signal states from a trusted source, using to our advantage that Eve may not hold the purification of the mixture.

There are several leading proof techniques for handling state preparation errors in a QKD protocol. The first major analysis was performed in [7]; however, the authors assumed pessimistically that Eve could amplify such noise to her benefit, so the technique was not robust over long distances against, e.g., coherent modulation errors. An improved technique was provided in the loss-tolerant protocol [8], which uses basis mismatch statistics to infer phase error rates that cannot be directly observed when state preparation is nonideal. However, the technique leaves ambiguous how to treat *mixed* signal states, a gap this work closes. Different extensions of the loss-tolerant protocol were considered in [9–11]; however, their focus was primarily on leaky sources, so treatment of mixed states was analogous to [8]. Another notable technique for characterizing security given pure qubit signal states is provided by [12]; however, their technique for generalizing to mixed signal states uses a suboptimal approach of averaging the key rates for each of the pure states in the mixture, which yields an equal or lower key rate than the key rate produced from the true average signal state. Lastly, an approach for finding a numerical lower bound on the Devetak-Winter secret key rate [13] for MDI-QKD protocols is provided by [14,15];

[*] bourassa@physics.utoronto.ca
[†] charles.lim@nus.edu.sg

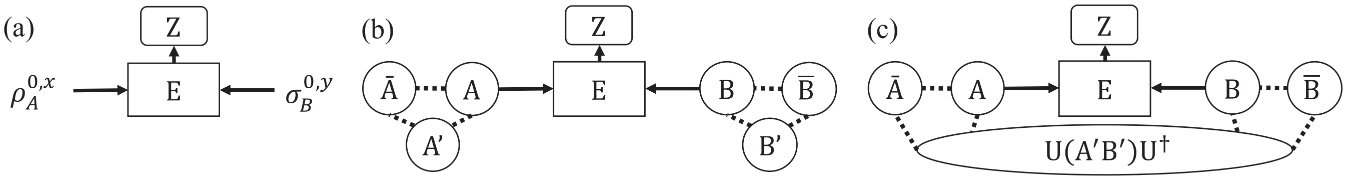

FIG. 1. (a) A real MDI QKD protocol: Alice and Bob send mixed states associated with bit (x, y) values to a central node controlled by Eve, who announces Z. (b) A virtual (purified) picture of sending the key generation states: Alice and Bob's mixed signal states are entangled with virtual qubits $\bar{A}\bar{B}$ which coherently store the bit values (x, y) until they are revealed by a computational-basis measurement. The AB systems are additionally purified by the $A'B'$ systems to account for trusted noise in the source. Only the A, B systems are sent to Eve. (c) An alternative virtual picture: All purifications are related by unitary operations applied to, in general, a joint purifying ancilla, yielding private states in $\bar{A}\bar{B}A'B'$. These "twisting" operations can optimally boost the secret key rate as they modify the phase error rates which Alice and Bob need to estimate. In (a)–(c), the signal states sent and the observed detection and bit error rates are the same.

their technique is in principle extendable to noisy state preparation. In our work, we take a conceptually simpler strategy of directly optimizing the key rate formula from [16], which uses the bit and phase errors of qubits in a virtual picture of the protocol.

In the case state preparation noise can be trusted and characterized, but perhaps not reduced; we provide here a simple analytical and numerical toolbox for calculating an optimal secret key rate. First, we provide a re-framing of the tilted four-state loss-tolerant protocol which provides a method for fixing Eve's degrees of freedom in the secret key rate [8,17,18]. However, as the signal states are mixed, the security also depends on how we treat the trusted noise in the signal state generation. Typically, the security of QKD is analyzed in terms of Alice and Bob's ability to virtually distill maximally entangled EPR pairs, since measurement of such pairs yields perfectly correlated keys, and by the monogamy of entanglement, the results cannot be correlated with anyone else, including Eve. However, it is known that a larger class of states known as private states [19–22] are fundamentally what is required to produce secret key. Formally, private states can be constructed from an EPR pair if Alice and Bob take ancillary shield systems they control, and apply a "twisting" unitary operation between the EPR pair and the shields, the condition being that this twisting leave unaffected the measurement results that generate secret key. Since twisting does not change the key, private states can then be understood as deflecting some of Eve's attack on the systems that generate key to the shield systems. See Fig. 1 for a diagram of this concept.

In our technique, we show that the mixing noise of the signal states can be treated in a virtual picture as being equivalent to Alice and Bob employing shield systems. Completely within this virtual picture, we can apply unitary twisting operations to the shields to decrease the phase errors of the protocol, increasing the secret key rate. We provide simple semidefinite programs to find the optimal twisting operations, yielding the optimal key rate under this framework. Semidefinite programming [23] has recently become a powerful tool for quantifying the asymptotic security of QKD protocols [14,15,24–29]. While private states have been of significant conceptual interest, as far as we are aware, this is the first application of private states in a practical QKD setting. Finally, we apply our technique to calculating fundamentally achievable key rates in an MDI QKD protocol with randomized modulation error in the state preparation procedure. We note that our technique is applicable to a wide class of MDI QKD protocols in which Alice and Bob each employ four qubit signal states that must not fall in the same plane of the Bloch sphere (which is easy to impose in practice), but which can be subject to general asymmetric preparation noise. Moreover, these signal states can be the single photon components of phase randomized coherent states in a decoy state protocol.

II. CHARACTERIZING EVE'S STATE

We consider an MDI QKD protocol in which Alice and Bob each prepare four mixed qubit signal states $\{\rho_A^{i,x}\}$ and $\{\sigma_B^{j,y}\}$, that they will send, respectively, with probabilities $p^{i,x}$ and $q^{j,y}$ to the central measurement node controlled by Eve. Alice and Bob can characterize their initial states by, e.g., performing tomography on their sources before the protocol, as in [17]. When Alice and Bob choose $(i, j) = (0, 0)$ these are the key generation states with (x, y) corresponding to their key bit values. All other combinations (i, j, x, y) correspond to test states used to constrain the phase errors. Following the security proof of the loss-tolerant protocol [8], we require that the sets of states $\{\rho_A^{i,x}\}$ and $\{\sigma_B^{j,y}\}$ each form a tetrahedron on the Bloch sphere, meaning the Bloch vectors cannot all lie in the same plane. In Appendix A, we provide steps for how to embed our technique within a decoy state protocol in the asymptotic limit of an infinite number of decoys. Note we are also assuming collective attacks, with an extension to coherent attacks available in [8].

As qubits, our signal states can be fully characterized with two orthonormal basis vectors, which we take to be the polarization states $|H\rangle$, $|V\rangle$:

$$\rho_A^{i,x}\sigma_B^{j,y} = \sum_{\substack{m,\,m',\\ n,\,n' = H}}^{V} c_{m,m'}^{i,x} d_{n,n'}^{j,y} \, |m, n\rangle \, \langle m', n'|_{A,B} \,. \qquad (1)$$

Under unitary evolution, each of these basis vectors evolves to a (subnormalized) state in Eve's possession as well as a classical announcement z, which we take to be pass or fail: $|m, n\rangle_{A,B} \rightarrow \sum_{z=P}^{F} |e_{m,n}^{z}\rangle_E \, |z\rangle_Z$. This process generalizes simply to multiple announcement events, such as which Bell state Eve claims to have detected.

The probabilities that Eve announces a round successfully passed conditioned on the signal states sent $p_{\text{det}}^{i,j,x,y} = p(z = P|i, j, x, y)$, provide constraints on the inner product of Eve's vectors $\langle e^P_{m',n'}|e^P_{m,n}\rangle_E$:

$$p_{\rm det}^{i,j,x,y} = p^{i,x}q^{j,y}\sum_{\substack{m,m',\\ n,n'=H}}^{V} c^{i,x}_{m,m'}d^{j,y}_{n,n'}\,\langle e^P_{m',n'}|e^P_{m,n}\rangle_E\,. \tag{2}$$

$p_{\rm det}^{i,j,x,y}$ are observable statistics in the protocol, and they can also be used to directly calculate some quantities required for the secret key rate formula, such as the detection probability in the key basis, $p_{\rm det}^{0,0} = \sum_{x,y} p_{\rm det}^{0,0,x,y}$, and the bit error rate $e_Z = (p_{\rm det}^{0,0,0,1} + p_{\rm det}^{0,0,1,0})/p_{\rm det}^{0,0}$, where we have taken $|\Phi^+\rangle$ to be the target Bell state that Alice and Bob wish to distill in a virtual picture we describe in the next section.

We see that Eq. (2) can be written compactly as

$$\vec{p}_{\rm det} = \hat{\gamma}\vec{e} \Rightarrow \vec{e} = \hat{\gamma}^{-1}\vec{p}_{\rm det}, \tag{3}$$

where $\vec{e}_s = \langle e^P_{m',n'}|e^P_{m,n}\rangle_E$, $s = 1, \ldots, 16$, is the vectorized form of the Gramian matrix of Eve's states associated with a passing announcement; $(\vec{p}_{\rm det})_t = p_{\rm det}^{i,j,x,y}$, $t = 1, \ldots, 16$, is a vector containing all the successful detection probabilities; and $\hat{\gamma}_{ts} = p^{i,x}q^{j,y}c^{i,x}_{m,m'}d^{j,y}_{n,n'}$ is a matrix dependent on the initial states from Eq. (1) used in the protocol. As long as $\hat{\gamma}^{-1}$ exists, then we can exactly solve for $\vec{e}$, which can then be used to calculate any objective function of $\langle e^P_{m',n'}|e^P_{m,n}\rangle_E$, including all the phase error rates in the six-state protocol key rate formula [1,8,16,30], even though we are only using four states, which were chosen to provide complete characterization of Eve's strategy. In Appendix B, we show that the invertibility of $\hat{\gamma}$ is equivalent to sending four states that form a tetrahedron on the Bloch sphere, as found in the original security proof of the tilted four-state loss-tolerant protocol [8]. Additionally, we provide a generalization of the proof technique to high-dimensional MDI QKD [27,31–34].

III. OPTIMAL CHOICE OF VIRTUAL PROTOCOL

Having characterized Eve's Gramian matrix entirely from observable parameters in the protocol, we now move to a virtual picture for the key generation signal states to calculate the remaining parameters of the secret key rate. In this virtual picture, which is depicted in Fig. 1, systems A, B from Eq. (1) are entangled with virtual qubits $\bar{A}$, $\bar{B}$ that Alice and Bob keep in their laboratory [8]. Importantly, since these signal states are mixed, we require additional purifying ancillary systems $A'B'$. We assume that the sources of noise are confined to Alice and Bob's labs, meaning Eve does not have access to manipulate $A'B'$. The mixedness of the signal states then results in an effective virtual shield Alice and Bob can use to minimize Eve's knowledge of the secret key.

The key generation states, $p^{0,x}q^{0,y}\rho_A^{0,x}\sigma_B^{0,y}$ can be purified to

$$|\zeta\rangle = \sum_{x,y}|x,y\rangle_{\bar{A}\bar{B}}\sum_{m,n=H}^{V}|\gamma^{x,y}_{m,n}\rangle_{A'B'}|m,n\rangle_{AB}, \tag{4}$$

where we have constraints from the states in Eq. (1):

$$\langle\gamma^{x,y}_{m',n'}|\gamma^{x,y}_{m,n}\rangle_{A'B'} = p^{0,x}q^{0,y}c^{0,x}_{m,m'}d^{0,y}_{n,n'}, \tag{5}$$

since to generate key, Alice and Bob measure $\bar{A}\bar{B}$ in the computational basis. The crucial point is that this purification is not unique [35], and so we have freedom to choose the virtual picture that yields the optimal key rate. Since Eve does not have access to $A'B'$, any purification will yield a suitable lower bound on the key rate, but we will show how to choose the optimal purification with simple semidefinite programs.

We can parametrize all purifications using twisting unitary operations [19–22] applied to the virtual ancillary systems in $|\zeta\rangle$:

$$U_{\bar{A}\bar{B}A'B'} = \sum_{x,y=0}^{1}|x,y\rangle\langle x,y|_{\bar{A}\bar{B}}\otimes U^{x,y}_{A'B'}. \tag{6}$$

Such an operation is entirely virtual, so it can be nonlocal in general and never needs to be executed in the real protocol. Twisting does not affect any of the real observed detection probabilities, which correspond to Alice and Bob first projecting $\bar{A}\bar{B}$ in the computational basis, as we show in Appendix C. Moreover, since only the A, B portion of $|\zeta\rangle$ evolves unitarily to E, Z, the twisting operation need not be fixed from the beginning of the protocol, and its choice can and should be informed by the statistics of the protocol. Such twisting operations applied to Bell states yield private states. We next show exactly how these twisting operations affect the secret key formula.

To quantify the security of the protocol, we employ the key rate formula from the six-state protocol [1,16,30], noting, however, that our protocol employs only four states:

$$\begin{aligned} R = p_{\rm det}^{0,0}\bigg(1 - h_2(e_Z) - e_Z h_2\bigg[\frac{1+(e_X-e_Y)/e_Z}{2}\bigg] \\ - (1-e_Z)h_2\bigg[\frac{1-(e_X+e_Y+e_Z)/2}{1-e_Z}\bigg]\bigg), \end{aligned} \tag{7}$$

where $h_2(\cdot)$ is the binary entropy function, and e_X and e_Y are the phase error rates of the virtual qubits $\bar{A}\bar{B}$ in the X and Y Pauli bases. These can be understood as the probability of the virtual qubits being projected into the incorrect Bell states given a passing announcement from Eve:

$$\begin{aligned} e_X &= \frac{\langle\Gamma|[(|\Psi^-\rangle\langle\Psi^-| + |\Phi^-\rangle\langle\Phi^-|)_{\bar{A}\bar{B}}\otimes|P\rangle\langle P|_Z]|\Gamma\rangle}{\langle\Gamma|(|P\rangle\langle P|_Z)|\Gamma\rangle}, \\ e_Y &= \frac{\langle\Gamma|[(|\Psi^+\rangle\langle\Psi^+| + |\Phi^-\rangle\langle\Phi^-|)_{\bar{A}\bar{B}}\otimes|P\rangle\langle P|_Z]|\Gamma\rangle}{\langle\Gamma|(|P\rangle\langle P|_Z)|\Gamma\rangle}. \end{aligned} \tag{8}$$

Here, $|\Gamma\rangle$ denotes the joint state between $\bar{A}\bar{B}A'B'EZ$ after the AB portion of twisted purified state $U_{\bar{A}\bar{B}A'B'}|\zeta\rangle$ is sent to Eve. Note $\langle\Gamma|(|P\rangle\langle P|_Z)|\Gamma\rangle = p_{\rm det}^{0,0}$. The six-state protocol key rate provides generally higher key rates than the Shor-Preskill key rate [36] because it takes into account correlations between the bit and phase error patterns.

Employing Eq. (4), the twisting operation in Eq. (6), and the unitary evolution $|m,n\rangle_{A,B} \rightarrow \sum_{z=P}^{F}|e^z_{m,n}\rangle_E|z\rangle_Z$, we find that $e_\pm = e_X \pm e_Y$ are linear functions with respect to the elements of Eve's Gramian matrix $\langle e^P_{m',n'}|e^P_{m,n}\rangle_E$, which are already known from Eq. (3). Additionally, these phase errors are linear with respect to matrix elements $\langle\gamma^{x',y'}_{m',n'}|U^{x',y'\dagger}_{A'B'}U^{x,y}_{A'B'}|\gamma^{x,y}_{m,n}\rangle_{A'B'}$, which are functions of the twisting operation we control. Since our task is to modify the twisting operation to boost the key rate, these elements form the optimization variables of our problem. We note these elements form the Gramian matrix of the twisted ancillary system states, which is a positive semidefinite (PSD) matrix by construction. They are constrained linearly by Eq. (5), since by construction when $(x, y) = (x', y')$, the twisting operations cancel to not affect the form of the real protocol signal states.

The additional benefit of choosing $e_\pm$ as our objective functions is that we overcome any nonlinear optimization introduced by $h_2(\cdot)$. We find that e_+ (e_-) only depends on $U_+ = U_{A'B'}^{0,0\,\dagger} U_{A'B'}^{1,1}$ ($U_- = U_{A'B'}^{0,1\,\dagger} U_{A'B'}^{1,0}$). Intuitively, we have such a dependence since e_- involves the Bell states that underwent a bit flip, so Alice and Bob's bit values will be (0,1) and (1,0), and only those twisting unitaries will be used. Similarly e_+ reflects Bell states that did not undergo a bit flip, so twisting will only involve (0,0) and (1,1). Since the unitaries $\{U_{A'B'}^{x,y}\}$ can be defined independently of each other, the optimizations of $e_\pm$ can be decoupled. Finally, since $h_2(x \leqslant 1/2)$ is monotonic, optimization of the arguments $e_\pm$ is sufficient.

Taking stock, we have two independent objective functions $e_\pm$, which are linear with respect to our optimization variables $\langle \gamma_{m',n'}^{x',y'} | U_{A'B'}^{x',y'\,\dagger} U_{A'B'}^{x,y} | \gamma_{m,n}^{x,y} \rangle_{A'B'}$, the elements of a PSD matrix subject to linear constraints. Thus, these optimization problems take the form of semidefinite programs which can be solved numerically on a standard laptop in a few seconds using packages for PYTHON [37,38] or MATLAB [39]. While previous literature on twisting operations had noted the opportunity for optimizing U [22], no explicit procedure was constructed. Here, we have closed this gap, increasing the practicality of utilizing a virtual twisting operation as a step in the security proof. In Appendix C, we provide complete details for framing the problem in terms of semidefinite programs.

A comment is in order regarding our ability to choose an optimistic and optimal purification to increase the key rate, as it is common in other QKD security proofs to assume Eve holds the purification and thus one might assume we need to choose the most pessimistic purification. Recall from Fig. 1 that the qubits from which the secret key is extracted are $\bar{A}\bar{B}$. These qubits are entangled with the signal states AB, and with $A'B'$, since the signal states are mixed. We are assuming that the sources are imperfect, but not malicious, so Eve does not have access to $A'B'$. This means that we are able to choose the virtual state of $A'B'$ in the most optimistic manner, which is equivalent to using $A'B'$ as a shield system upon which we can apply twisting operations to yield private states with $\bar{A}\bar{B}$. After the initial states in AB are sent to Eve, who also holds system E, Eve then holds a partial purification of the key systems $\bar{A}\bar{B}$ as well, but she still does not hold the entire purification since she does not have $A'B'$. Indeed, for a general protocol we would need to determine the most pessimistic state Eve could hold, which would correspond to finding the most pessimistic form of her Gramian with respect to the secret key rate; however, as we are using the tilted four-state protocol, we have from Eq. (3) that Eve's Gramian is fixed, so there is no room to modify the parameters of the secret key rate that depend on Eve. That is, we do not need to take the most pessimistic partial purification that Eve can hold, because we have already uniquely specified her Gramian. Thus, the only remaining free parameters come from the state of the shield system, which we have the benefit of treating optimally by applying the twisting operation in Eq. (6). Picking the optimal purification of virtual systems Eve cannot access has been used to advantage in QKD security proofs before, as in choosing the state for the fictitious quantum coin in [7].

IV. KEY RATE RESULTS

The only requirement for applying our technique is that Alice and Bob's initial qubit signal states cannot fall in the same plane of the Bloch sphere, which is easy to satisfy in practice. Otherwise, our technique can handle quite general noisy state preparation: Alice and Bob need not prepare the same sets of states; they can send states with different probabilities; and, the noise channel applied to each state can be dependent on the state.

As a study of fundamentally achievable key rates, we consider the following two-parameter (δ, p) model for the initial states. We suppose Alice and Bob attempt to prepare the states $\{|H\rangle, |V\rangle, {}^{|H\rangle+|V\rangle}/_{\sqrt{2}}, {}^{|H\rangle-i|V\rangle}/_{\sqrt{2}}\}$; however, each state is subject to a modulation error which we treat as a random variable. Given a state-dependent distribution for the modulation error on the surface of the Bloch sphere, the resulting average states can be treated as having a coherent modulation error, i.e., a constant offset angle from the ideal state parametrized by δ, as well as a depolarization noise parametrized by p, which introduces incoherent mixing to the states, shortening the Bloch vector. For exact definitions of the signal states, see Appendix E. For the case $p = 0$, we expect no improvement in our key rate over the standard loss-tolerant protocol, since no mixing implies no virtual ancillary shield system.

In Fig. 2, we plot the asymptotic key rate found using our technique as a function of distance for various pairs (δ, p). For comparison with the key rate produced with our optimization, we plot the key rate calculated using a suboptimal purification, which was constructed by diagonalizing Alice and Bob's signal states and having $A'B'$ index the eigenvalues in decreasing order. We find that our technique provides a modest increase over the "naive" purification, our technique's advantages being most significant as the depolarizing noise gets stronger (making the initial states more mixed), and at longer distances when the untrusted channel noises (loss and dark counts) accrue. Additionally, we see a better key rate can be produced by reducing state preparation noise; however, once one has improved as best possible, our technique provides an optimized key rate given that level of noise. That is, our technique provides confidence that one has optimized over all possible ancillary states of the purification that are consistent with the protocol statistics without worry that one has chosen a pessimistic virtual picture.

V. CONCLUSION

We have presented an extension of the proof technique from [8] to quantify the security of MDI QKD protocols that employ general noisy qubit signal states. We first reframed the analytical technique used for characterizing the parameters in the secret key rate that depend on Eve's system, noting that this approach lends itself clearly to a generalization for higher-dimensional signal states. Next, we observed that employing trusted but mixed signal states means Alice and Bob have not a single but a set of virtual pictures they can use to analyze security in their protocol; we observed this was equivalent to

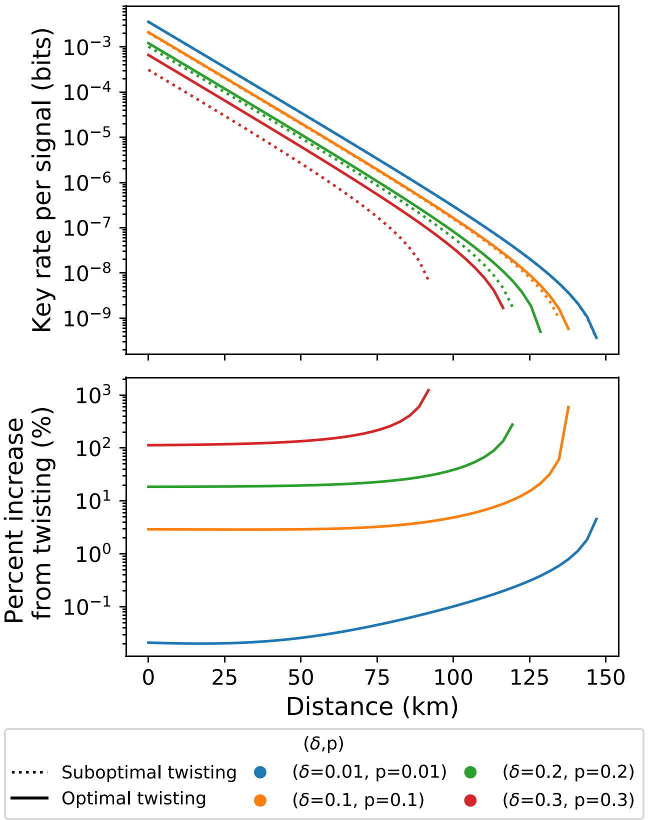

FIG. 2. (Top) Key rate vs Alice-Charlie distance for various values of modulation error and depolarizing noise (δ, p) (same color indicates same model parameters). The dotted lines are the results from a suboptimal purification, while solid lines indicate our optimized key rates over all virtual twisting operations. (Bottom) For each pair (δ, p), the percentage increase offered by optimizing over twisting operations. We assume a single photon source, symmetric distances from Alice and Bob to Charlie, and a Bell state detection scheme similar to [3], with overall detection efficiency of 50%, a dark count probability of 10^{-5} per pulse per detector, loss in fiber of 0.2 dB/km, and error correction efficiency of 1.

Alice and Bob employing a *virtual* shield system onto which they can apply *virtual* twisting operations to minimize Eve's knowledge of the key [19,20]. Finally, we provided a simple numerical technique leveraging semidefinite programming to optimize over all twisting operations to optimize the six-state protocol secret key rate formula, examining the implications for state preparation subject to random modulation error.

ACKNOWLEDGMENTS

We thank Ilan Tzitrin, Wenyuan Wang, Thomas van Himbeeck, Emilien Lavie, and Koon Tong Goh for useful discussions. We thank Natural Sciences and Engineering Research Council of Canada (NSERC), CFI, ORF, MITACS, Royal Bank of Canada, Huawei Technology Canada, NUS-UT collaborative Ph.D cluster funding, and HKU start-up grants for funding support. J.E.B. is supported by an Ontario Graduate Scholarship and the Lachlan Gilchrist Fellowship. I.W.P. is supported by National Research Foundation and the Ministry of Education, Singapore, under the Research Centres of Excellence programme. C.C.W.L. acknowledges support from the National Research Foundation (NRF) Singapore, under its NRF Fellowship programme (NRFF11-2019-0001) and Quantum Engineering Programme (QEP-P2).

APPENDIX A: EMBEDDING OUR TECHNIQUE WITHIN A DECOY STATE PROTOCOL

In the main text, we showed that, using two independent semidefinite programs (SDP), we can optimize the secret key rate of a loss-tolerant protocol where the signal states are two-dimensional mixed states. In the context of quantum key distribution, those states are normally encoded in the mode (such as polarization or time bin) of single photons. However, in practice, many protocols employ weak coherent pulses with decoy states [5,6]. In this section, we outline how we can embed our technique in a loss-tolerant protocol which uses decoy states. In this work, we will consider a protocol with an infinite number of decoy states as in [8]. We leave the case where a finite number of decoy states are used for future work. Like in the main text, we work in the asymptotic limit where we can ignore finite key effects.

Recall from the main text that to use our technique, we need two pieces of information: Before sending optical signals to Eve, we need the density matrices of the qubit signal states (the single photon component), and then once the optical signals are sent, we need the probability of successful Bell state measurement given the states that Alice and Bob chose. As such, to apply our technique when phase-randomized weak coherent pulses are used together with a decoy state protocol, we need to calculate the state of the single photon component of the optical signal as well as the probability of successful Bell state measurement given that Alice and Bob send the corresponding single photon signals. In the literature, that conditional probability is often referred to as the single photon yield, denoted by $Y_{11}^{i,j,x,y}$.

To obtain the single photon component of the signals, we can simply project the coherent signal states to their single photon components. Suppose that Alice and Bob prepare the phase-randomized coherent state $\tilde{\rho}_A^{i,x,\mu_A}$ and $\tilde{\sigma}_B^{j,y,\mu_B}$ with intensity μ_A and μ_B, respectively, the single photon components of those states can be easily obtained by performing the following projection and then normalizing the resulting states:

$$\begin{aligned}
\rho_A^{i,x} &= \frac{\left(|0\rangle\langle 0|_{H_A}\otimes|1\rangle\langle 1|_{V_A}+|1\rangle\langle 1|_{H_A}\otimes|0\rangle\langle 0|_{V_A}\right)\tilde{\rho}_A^{i,x,\mu_A}\left(|0\rangle\langle 0|_{H_A}\otimes|1\rangle\langle 1|_{V_A}+|1\rangle\langle 1|_{H_A}\otimes|0\rangle\langle 0|_{V_A}\right)}{e^{-\mu_A}\mu_A},\\
\sigma_B^{j,y} &= \frac{\left(|0\rangle\langle 0|_{H_B}\otimes|1\rangle\langle 1|_{V_B}+|1\rangle\langle 1|_{H_B}\otimes|0\rangle\langle 0|_{V_B}\right)\tilde{\sigma}_B^{j,y,\mu_B}\left(|0\rangle\langle 0|_{H_B}\otimes|1\rangle\langle 1|_{V_B}+|1\rangle\langle 1|_{H_B}\otimes|0\rangle\langle 0|_{V_B}\right)}{e^{-\mu_B}\mu_B},
\end{aligned} \tag{A1}$$

where $|0\rangle_m$ and $|1\rangle_m$ are the vacuum and single photon states in mode m, respectively.

Hence, it is important that we characterize the sources of each legitimate party before performing the protocol. Ideally, this should be done by performing tomography on the signal states $\tilde{\rho}_A^{i,x,\mu_A}$ and $\tilde{\sigma}_B^{j,y,\mu_B}$. Alternatively, one can have a model for the source, taking into account the finite precision and randomness in the modulation of the signal states. Once we have the single photon component of the signal states [i.e., $\rho_A^{i,x}$ and $\sigma_B^{j,y}$ in Eq. (1) of the main text], we can use them to construct the $\hat{\gamma}$ matrix in Eq. (3) of the main text, and to impose the constraints on the ancillary systems $A'B'$ as described in Eq. (5) of the main text.

On the other hand, from the parameter estimation step of the protocol, we can estimate the gain $Q_{\mu_A,\mu_B}^{i,j,x,y}$ for each choice of states (i, x) and (j, y) and intensities μ_A, μ_B. When using an infinite number of decoy states, Alice and Bob can determine the values of the single photon yield $Y_{11}^{i,j,x,y}$ exactly for all i, j, x, y. Once the values of $Y_{11}^{i,j,x,y}$ are obtained, we can replace the $p_{\text{det}}^{i,j,x,y}$ with $Y_{11}^{i,j,x,y}$ in Eq. (3) of the main text and then proceed with our method.

APPENDIX B: RELATIONSHIP BETWEEN THE INVERTIBILITY OF $\hat{\gamma}$ AND THE STATES IN THE BLOCH SPHERE FORMING A TETRAHEDRON

In the main text, we demonstrated that we could solve for the elements of Eve's Gramian matrix using the following equation:

$$\vec{p}_{\text{det}} = \hat{\gamma}\vec{e} \Rightarrow \vec{e} = \hat{\gamma}^{-1}\vec{p}_{\text{det}}, \tag{B1}$$

where $\vec{e}_s = \langle e^P_{m',n'}|e^P_{m,n}\rangle_E$ are the elements of the vectorized form of the Gramian matrix of Eve's states associated with a passing announcement. $(\vec{p}_{\text{det}})_t = p_{\text{det}}^{i,j,x,y}$ form a vector containing all the successful detection probabilities, and $\hat{\gamma}_{ts} = p^{i,x}q^{j,y}c^{i,x}_{m,m'}d^{j,y}_{n,n'}$ form a matrix dependent on the initial states used in the protocol which were taken to be

$$\rho_A^{i,x}\sigma_B^{j,y} = \sum_{\substack{m,m',\\ n,n'=H}}^{V} c^{i,x}_{m,m'}d^{j,y}_{n,n'}\,|m,n\rangle\,\langle m',n'|_{A,B}\,. \tag{B2}$$

Here we show that the invertibility of $\hat{\gamma}$ is equivalent to the condition in the loss-tolerant protocol [8] that Alice and Bob each need to choose four signal states that form a tetrahedron in the Bloch sphere, as shown in Fig. 3.

We begin by noting that

$$\hat{\gamma}_{ts} = p^{i,x}q^{j,y}c^{i,x}_{m,m'}d^{j,y}_{n,n'} = p^{i,x}q^{j,y}\,\langle m,n|\rho_A^{i,x}\sigma_B^{j,y}|m',n'\rangle_{A,B}\,, \tag{B3}$$

meaning we can always choose the basis ordering of $\hat{\gamma}$ so that its rows are $\text{vec}(p^{i,x}\rho_A^{i,x})^T \otimes \text{vec}(q^{j,y}\sigma_B^{j,y})^T$, the tensor product of the vectorized forms of the probability-weighted signal states. Invertibility of $\hat{\gamma}$ is equivalent to showing its rows are linearly independent. Since all the row vectors have tensor product form, we just need to show that the $\{\text{vec}(p^{i,x}\rho_A^{i,x})\}$ and the $\{\text{vec}(q^{j,y}\sigma_B^{j,y})\}$ each form sets of linearly independent vectors.

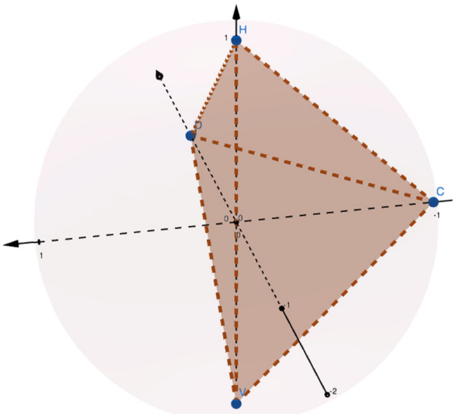

FIG. 3. A tetrahedron in the Bloch sphere representing qubits encoded in horizontal, vertical, diagonal, and clockwise circularly polarized single photons. A tetrahedron is formed so long as the states don't all fall in the same plane.

Next, we recall that four states forming a tetrahedron in the Bloch sphere is equivalent to them having linearly independent Stokes vectors. Let's focus on Alice's states, since the states are qubits, they can be expressed in terms of Stokes parameters:

$$p^{i,x}\rho_A^{i,x} = \frac{1}{2}\sum_{r=0}^{3} P_r^{i,x}\sigma_r, \tag{B4}$$

where σ_0 is the identity and σ_r, $r = 1, 2, 3$ are the Pauli matrices, while $P_r^{i,x} = p^{i,x}\text{Tr}(\sigma_r\rho_A^{i,x})$ form the elements of the Stokes vector $\vec{P}^{i,x}$ for that state. Thus,

$$\text{vec}\big(p^{i,x}\rho_A^{i,x}\big) = \frac{1}{2}\sum_{r=0}^{3} P_r^{i,x}\text{vec}(\sigma_r). \tag{B5}$$

It is easy to show that $\text{vec}(\sigma_r)^T\text{vec}(\sigma_{r'}) = \delta_{rr'}$, which means the inner product of any two $\text{vec}(p^{i,x}\rho_A^{i,x})$ is related to the inner product of the Stokes vectors by a constant factor:

$$\text{vec}^T\left(p^{i,x}\rho_A^{i,x}\right)\text{vec}\big(p^{i',x'}\rho_A^{x'i'}\big) = \frac{1}{2}\sum_{r=0}^{3} P_r^{i,x}P_r^{x',i'} = \frac{1}{2}\vec{P}^{i,x}\cdot\vec{P}^{x',i'}. \tag{B6}$$

Thus, since the inner product structure of the rows of $\hat{\gamma}$ is identical to that of the Stokes vectors up to a factor, the linear independence of the Stokes vectors is equivalent to the invertibility of $\hat{\gamma}$.

Generalization to high-dimensional protocols. Having reframed the security proof from the loss-tolerant protocol in this form, we observe that the matrix inversion Eq. (B1) can also generalize straightforwardly to MDI QKD protocols employing discrete-variable high-dimensional degrees of freedom, such as those employing orbital angular momentum [31–33] or time-bin encodings [27,34].

For Alice and Bob each sending d-dimensional systems, Eve's Gramian matrix will have d^4 elements that we can flatten to a vector $\vec{e}$. Thus, Alice and Bob can each prepare d^2 states within the d-dimensional space, that will in turn yield d^4 observable detection probabilities that we can store in the vector $\vec{p}_{\text{det}}$. The basis ordering of $\hat{\gamma}$ can be chosen so that its rows are the tensor product of the vectorized forms of the probability-weighted signal states $\mathrm{vec}(p^{i,x}\rho_A^{i,x})^T \otimes \mathrm{vec}(q^{j,y}\sigma_B^{j,y})^T$. Since invertibility of $\hat{\gamma}$ is equivalent to its rows being linearly independent, this provides a clear minimum requirement for state preparation, namely that the vectorized forms of the states which Alice and Bob use in the protocol must be linearly independent.

If they can prepare their states so that the d^4 rows of $\hat{\gamma}$ are linearly independent, then they will be able to satisfy Eq. (B1). This condition of linear independence for the vectorized density matrices is a much less stringent condition to satisfy for high-dimensional protocols than, e.g., employing mutually unbiased bases [40,41], while still allowing for complete characterization of the parameters in the high-dimensional secret key rate formula that are dependent on Eve's system [42].

APPENDIX C: THE VIRTUAL PICTURE AND OPTIMIZATION OF THE KEY RATE WITH SEMIDEFINITE PROGRAMMING

In the main text, we provided an overview of how to determine the optimal virtual picture for our protocol using a virtual twisting operation [19] and semidefinite programming. Here we provide the full mathematical details of our analytical and numerical techniques.

1. Moving to a virtual picture

Given states of the form in Eq. (1), we can define a virtual purified picture for the key generation states:

$$\{p^{0,x}\rho_A^{0,x}q^{0,y}\sigma_B^{0,y}\} \to |\zeta\rangle_{\bar{A}\bar{B}A'B'AB} = \sum_{x,y}|x,y\rangle_{\bar{A}\bar{B}}\sum_{m,n=H}^{V}\left|\gamma_{m,n}^{x,y}\right\rangle_{A'B'}|m,n\rangle_{AB}, \tag{C1}$$

where we know from the initial states that $\mathrm{Tr}_{\bar{A}\bar{B}A'B'}(|x,y\rangle\langle x,y|_{\bar{A}\bar{B}}\,|\zeta\rangle\langle\zeta|_{\bar{A}\bar{B}A'B'AB}) = p^{0,x}\rho_A^{0,x}q^{0,y}\sigma_B^{0,y}$, which fixes the constraint:

$$\left\langle\gamma_{m',n'}^{x,y}\middle|\gamma_{m,n}^{x,y}\right\rangle_{A'B'} \equiv p^{0,x}q^{0,y}c_{m,m'}^{0,x}d_{n,n'}^{0,y}, \tag{C2}$$

where these also correspond to a subset of the matrix elements from Eq. (B3) for the key generation states.

The constraint in Eq. (C2) on the ancillary systems is not sufficient to identically fix the purification. Rather, any unitary twisting operation of the form,

$$U_{\bar{A}\bar{B}A'B'} = \sum_{x,y=0}^{1}|x,y\rangle\langle x,y|_{\bar{A}\bar{B}}\otimes U_{A'B'}^{x,y}, \tag{C3}$$

preserves the real signal states, since

$$\begin{aligned}&\mathrm{Tr}_{\bar{A}\bar{B}A'B'}(|x,y\rangle\langle x,y|_{\bar{A}\bar{B}}\,U_{\bar{A}\bar{B}A'B'}\,|\zeta\rangle\langle\zeta|_{\bar{A}\bar{B}A'B'AB}\,U_{\bar{A}\bar{B}A'B'}^{\dagger})\\ &\quad= \mathrm{Tr}_{\bar{A}\bar{B}A'B'}\big(|x,y\rangle\langle x,y|_{\bar{A}\bar{B}}\,U_{A'B'}^{x,y}\,|\zeta\rangle\langle\zeta|_{\bar{A}\bar{B}A'B'AB}\,U_{A'B'}^{x,y\,\dagger}\big)\\ &\quad= p^{0,x}\rho_A^{0,x}q^{0,y}\sigma_B^{0,y}.\end{aligned} \tag{C4}$$

Since the produced signal states are independent of this twisting operation, it cannot affect any of the real detection probabilities observed in the execution of the protocol which depend only on the A, B systems. Thus, the characterization of Eve's Gramian matrix elements $\langle e_{m',n'}^P|e_{m,n}^P\rangle_E$ is independent of the twisting operation. Moreover, since the produced signal states are independent of the twisting operation it never needs to actually be implemented in the real protocol, and can remain a useful virtual analytical tool in the characterization of security after the real signal exchange. Finally, the twisting operation can be chosen after the detection statistics are produced, which gives Alice and Bob the power to adjust their virtual strategy based on what they observe.

The major change when adding the twisting operation is the *definition* of the phase error rates, i.e., how we use the information about Eve's state from $\vec{e}_s$. Consider the joint state between Alice, Bob, their ancillae, and Eve in the purified picture after they apply a twisting operation and send the systems A, B to Eve:

$$\begin{aligned}U_{\bar{A}\bar{B}A'B'}\,|\zeta\rangle_{\bar{A}\bar{B}A'B'AB} &\to |\Gamma(U)\rangle_{\bar{A}\bar{B}A'B'EZ}\\ &= \sum_{x,y}|x,y\rangle_{\bar{A}\bar{B}}\sum_{m,n=H}^{V}U_{A'B'}^{x,y}\left|\gamma_{m,n}^{x,y}\right\rangle_{A'B'}\\ &\quad\times\sum_{z=P}^{F}\left|e_{m,n}^{z}\right\rangle_E|z\rangle_Z.\end{aligned} \tag{C5}$$

Thus, taking the target virtual Bell state to be $|\Phi^+\rangle\langle\Phi^+|_{\bar{A}\bar{B}}$, the phase error rates now become the *twisted phase error rates*:

$$\begin{aligned}e_X(U) &= \frac{1}{p_{\mathrm{det}}^{0,0}}\langle\Gamma(U)|[(|\Psi^-\rangle\langle\Psi^-| + |\Phi^-\rangle\langle\Phi^-|)_{\bar{A}\bar{B}}\otimes|P\rangle\langle P|_Z]|\Gamma(U)\rangle_{\bar{A}\bar{B}A'B'EZ}\\ &= \frac{1}{2} - \frac{1}{p_{\mathrm{det}}^{0,0}}\sum_{m,n,m',n'}\mathrm{Re}\big[\big(\big\langle\gamma_{m',n'}^{0,0}\big|U_{A'B'}^{0,0\,\dagger}U_{A'B'}^{1,1}\big|\gamma_{m,n}^{1,1}\big\rangle_{A'B'} + \big\langle\gamma_{m',n'}^{0,1}\big|U_{A'B'}^{0,1\,\dagger}U_{A'B'}^{1,0}\big|\gamma_{m,n}^{1,0}\big\rangle_{A'B'}\big)\big\langle e_{m',n'}^P\big|e_{m,n}^P\big\rangle_E\big],\\ e_Y(U) &= \frac{1}{p_{\mathrm{det}}^{0,0}}\langle\Gamma(U)|[(|\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+|)_{\bar{A}\bar{B}}\otimes|P\rangle\langle P|_Z]|\Gamma(U)\rangle_{\bar{A}\bar{B}A'B'EZ}\\ &= \frac{1}{2} - \frac{1}{p_{\mathrm{det}}^{0,0}}\sum_{m,n,m',n'}\mathrm{Re}\big[\big(\big\langle\gamma_{m',n'}^{0,0}\big|U_{A'B'}^{0,0\,\dagger}U_{A'B'}^{1,1}\big|\gamma_{m,n}^{1,1}\big\rangle_{A'B'} - \big\langle\gamma_{m',n'}^{0,1}\big|U_{A'B'}^{0,1\,\dagger}U_{A'B'}^{1,0}\big|\gamma_{m,n}^{1,0}\big\rangle_{A'B'}\big)\big\langle e_{m',n'}^P\big|e_{m,n}^P\big\rangle_E\big].\end{aligned} \tag{C6}$$

In general, there exists an optimal purification provided by some $U_{\bar{A}\bar{B}A'B'}$ such that the key rate is maximized. As it is unlikely to choose this optimal purification at random when constructing the problem, we expect the calculation of $e_X(U)$ and $e_Y(U)$ to benefit from an optimization over $U_{\bar{A}\bar{B}A'B'}$, and thus, in general, for the key rate to increase by employing an optimal twisted phase error rate.

Remark. Employing noisy, i.e., mixed, signal states in the protocol begs the question of how best to purify the states in the virtual protocol when defining the phase error rate. This necessarily leads to the definition of a twisted phase error rate and a search for an optimal twisting operation with respect to the key rate. The twisting operation is entirely in the virtual picture, so the optimal $U_{\bar{A}\bar{B}A'B'}$ can and should be computed after the exchange of signals and observation of detection probabilities and bit error rates. Although the form of the optimal twisting operation can even be nonlocal across $\bar{A}\bar{B}$ in practice, it does not need to be implemented in the real protocol, so its locality does not matter.

It is worth emphasizing that employing a twisting operation is an optional step in the security proof. Indeed, any phase error rates of the form Eq. (C6) will supply a suitable lower bound on the key rate, since Eve does not have control over the $A'B'$ systems. Nonetheless, optimizing over the twisting operation will in general boost the key rate, safeguarding against a poor, overly pessimistic initial choice of purification.

We next demonstrate how optimizing the phase error rates over twisting operations can be framed as two semidefinite programs, closing a gap in the previous literature on twisting operations [22].

2. Semidefinite programs for evaluating the six-state key rate

A general optimization problem is of the form [23],

$$\begin{aligned}&\texttt{minimize}\ f_0(\mathbf{x})\\&\texttt{s.t.}\quad f_i(\mathbf{x}) \geqslant b_i,\ i = 1, \ldots, m,\end{aligned}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ are the variables over which we optimize; $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function; $f_i : \mathbb{R}^n \to \mathbb{R}$ are the constraint functions; and, b_i are the constraint bounds. An optimal solution, $\mathbf{x}^\star$ would mean that for all $\mathbf{z}$ such that $f_i(\mathbf{z}) \geqslant b_i$ then $f_0(\mathbf{z}) \geqslant f_0(\mathbf{x}^\star)$.

Semidefinite programs (SDPs) are a class of convex optimization problems with linear objective and constraint functions over a cone of positive semidefinite (PSD) matrices [23]. That is, the optimization variables $\mathbf{x}$ form the elements of a matrix with non-negative eigenvalues, and $f_i(x) = \mathbf{c}_i \cdot \mathbf{x}$. They have become an incredibly versatile tool for QKD security proofs in recent years [14,15,24–27].

a. Objective functions

At first glance, the optimization problem required for the six-state key rate in Eq. (7) of the main text looks daunting. It appears we have two quantities to optimize with the twisting operation, $e_X(U)$ and $e_Y(U)$, appearing in a nonlinear function due to the binary entropy. However, consider a simple change of variable so that the two unknowns are given by

$$e_-(U) = (e_X - e_Y)(U),\ e_+(U) = (e_X + e_Y)(U). \tag{C7}$$

These remain linear objective functions of the only free variables in the problem $\{\langle\gamma^{x',y'}_{m',n'}|U^{x',y'\dagger}_{A'B'} U^{x,y}_{A'B'}|\gamma^{x,y}_{m,n}\rangle_{A'B'}\}$:

$$e_+(U) = 1 - \frac{2}{p^{0,0}_{\text{det}}} \sum_{m,n,m',n'} \text{Re}\big(\big\langle\gamma^{0,0}_{m',n'}\big|U^{0,0\dagger}_{A'B'} U^{1,1}_{A'B'}\big|\gamma^{1,1}_{m,n}\big\rangle_{A'B'} \big\langle e^P_{m',n'}\big|e^P_{m,n}\big\rangle_E\big) = e_+\big(U^{0,0\dagger}_{A'B'} U^{1,1}_{A'B'}\big), \tag{C8}$$

and

$$e_-(U) = -\frac{2}{p^{0,0}_{\text{det}}} \sum_{m,n,m',n'} \text{Re}\big(\big\langle\gamma^{0,1}_{m',n'}\big|U^{0,1\dagger}_{A'B'} U^{1,0}_{A'B'}\big|\gamma^{1,0}_{m,n}\big\rangle_{A'B'} \big\langle e^P_{m',n'}\big|e^P_{m,n}\big\rangle_E\big) = e_-\big(U^{0,1\dagger}_{A'B'} U^{1,0}_{A'B'}\big). \tag{C9}$$

The only free parameters over which we can optimize are the twisting unitaries, $\{U^{0,1}_{A'B'}, U^{1,0}_{A'B'}, U^{0,0}_{A'B'}, U^{1,1}_{A'B'}\}$, where each of the four unitaries can be defined independently of the others. Here, we have found that the two objective functions in the key rate $e_\pm(U)$ are functions of independent variables: $e_+(U)$ only depends on $U_+ = U^{0,0\dagger}_{A'B'} U^{1,1}_{A'B'}$ and $e_-(U)$ only depends on $U_- = U^{0,1\dagger}_{A'B'} U^{1,0}_{A'B'}$. This is very good, since it means the difficult task of nonlinear optimization of the six-state key rate formula can be avoided. Using the monotonicity of the binary entropy, we can directly optimize $e_\pm(U_\pm)$ within the binary entropy functions:

$$\begin{aligned}R &= \max_U p^{0,0}_{\text{det}}\left(1 - h_2(e_Z) - e_Z h_2\left[\frac{1 + e_-(U)/e_Z}{2}\right] - (1 - e_Z) h_2\left[\frac{1 - [e_+(U) + e_Z]/2}{1 - e_Z}\right]\right)\\ &= p^{0,0}_{\text{det}}\left(1 - h_2(e_Z) - e_Z h_2\left[\frac{1 + \max_{U_-} e_-(U_-)/e_Z}{2}\right] - (1 - e_Z) h_2\left[\frac{1 - [\min_{U_+} e_+(U_+) + e_Z]/2}{1 - e_Z}\right]\right),\end{aligned} \tag{C10}$$

with the extra conditions $0 \leqslant e_-(U_-) \leqslant e_Z$ and $e_Z \leqslant e_+(U_+) \leqslant 1$ so that the arguments of the binary entropy functions remain between 0 and 1.

b. Two independent semidefinite programs

We recall that $\langle e^P_{m',n'}|e^P_{m,n}\rangle_E$ are known from Eq. (3), while $\langle\gamma^{x',y'}_{m',n'}|U^{x',y'\dagger}_{A'B'} U^{x,y}_{A'B'}|\gamma^{x,y}_{m,n}\rangle_{A'B'}$ are the optimization variables. This leads to two independent semidefinite programs.

(i) For the linear objective function $e_-(U_-)$, we note the optimization variables,

$$\big\langle\gamma^{x,(x+1 \bmod 2)}_{m',n'}\big|U^{x,(x+1 \bmod 2)\dagger}_{A'B'} U^{y,(y+1 \bmod 2)}_{A'B'}\big|\gamma^{y,(y+1 \bmod 2)}_{m,n}\big\rangle_{A'B'}, \tag{C11}$$

form the 8×8 positive semidefinite Gram matrix for the vectors $\{U^{0,1}_{A'B'}\,|\gamma^{0,1}_{m,n}\rangle_{A'B'}, U^{1,0}_{A'B'}\,|\gamma^{1,0}_{m,n}\rangle_{A'B'}\}$, subject to the eight linear constraints from Eq. (C2):

$$\left\langle\gamma^{x,(x+1 \bmod 2)}_{m',n'}\middle|U^{x,(x+1 \bmod 2)\dagger}_{A'B'}U^{x,(x+1 \bmod 2)}_{A'B'}\middle|\gamma^{x(x+1 \bmod 2)}_{m,n}\right\rangle_{A'B'} = \left\langle\gamma^{x,(x+1 \bmod 2)}_{m',n'}\middle|\gamma^{x,(x+1 \bmod 2)}_{m,n}\right\rangle_{A'B'}. \tag{C12}$$

The optimization is additionally constrained by $0 \leqslant e_-(U_-) \leqslant e_Z$.

(ii) For the linear objective function $e_+(U_+)$, we note the optimization variables,

$$\left\langle\gamma^{x,x}_{m',n'}\middle|U^{x,x\dagger}_{A'B'}U^{y,y}_{A'B'}\middle|\gamma^{y,y}_{m,n}\right\rangle_{A'B'}, \tag{C13}$$

form the 8×8 PSD Gram matrix for the vectors $\{U^{0,0}_{A'B'}\,|\gamma^{0,0}_{m,n}\rangle_{A'B'}, U^{1,1}_{A'B'}\,|\gamma^{1,1}_{m,n}\rangle_{A'B'}\}$, subject to the eight linear constraints from Eq. (C2):

$$\left\langle\gamma^{x,x}_{m',n'}\middle|U^{x,x\dagger}_{A'B'}U^{x,x}_{A'B'}\middle|\gamma^{x,x}_{m,n}\right\rangle_{A'B'} = \left\langle\gamma^{x,x}_{m',n'}\middle|\gamma^{x,x}_{m,n}\right\rangle_{A'B'}. \tag{C14}$$

The optimization is additionally constrained by $e_Z \leqslant e_+(U_+) \leqslant 1$.

With that, we have two independent semidefinite programs which can be used to optimize the six-state key rate formula. In Appendix D, we provide a pseudocode overview of our numerical technique for calculating the key rate.

APPENDIX D: PSEUDOCODE FOR KEY RATE CALCULATION

Here we present a sketch of our numerical implementation for calculating key rates. For the semidefinite programs we employed CVXPY [37,38], a convex optimization library for PYTHON. All codes are available upon request.

APPENDIX E: (δ, p) MODEL FOR SIGNAL STATES

We consider the following two-parameter (δ, p) model for the initial states which Alice and Bob prepare:

$$\begin{aligned}
\rho^{0,0}_A &= \sigma^{0,0}_B = (1-p)\left|\xi^\delta_{00}\right\rangle\left\langle\xi^\delta_{00}\right| + {}^p\!/\!_2\mathbb{1},\\
\rho^{0,1}_A &= \sigma^{0,1}_B = (1-p)\left|\xi^\delta_{01}\right\rangle\left\langle\xi^\delta_{01}\right| + {}^p\!/\!_2\mathbb{1},\\
\rho^{1,0}_A &= \sigma^{1,0}_B = (1-p)\left|\xi^\delta_{10}\right\rangle\left\langle\xi^\delta_{10}\right| + {}^p\!/\!_2\mathbb{1},\\
\rho^{1,1}_A &= \sigma^{1,1}_B = (1-p)\left|\xi^\delta_{11}\right\rangle\left\langle\xi^\delta_{11}\right| + {}^p\!/\!_2\mathbb{1},
\end{aligned} \tag{E1}$$

where the states $|\xi^\delta\rangle$ are of the form,

$$\begin{aligned}
&\left|\xi^\delta_{00}\right\rangle = |H\rangle, \quad \left|\xi^\delta_{01}\right\rangle = -\sin\frac{\delta}{2}|H\rangle + \cos\frac{\delta}{2}|V\rangle,\\
&\left|\xi^\delta_{10}\right\rangle = \cos\frac{\pi+\delta}{4}|H\rangle + \sin\frac{\pi+\delta}{4}|V\rangle,\\
&\left|\xi^\delta_{11}\right\rangle = \cos\frac{-\pi+\delta}{4}|H\rangle + i\sin\frac{-\pi+\delta}{4}|V\rangle.
\end{aligned} \tag{E2}$$

Algorithm 1: Key rate function

```
function KEYRATE($\rho_A, \sigma_B, p_A, q_B, p_{\rm dark}, \eta, l$)
    # $\rho_A$ and $\sigma_B$ are arrays containing Alice and Bob's four density
      matrices, s.t. $\rho_A[i,x] = \rho^{i,x}_A$, $\sigma_B[j,y] = \sigma^{j,y}_B$
    # $p_A$ and $q_B$ are lists of the probabilities for sending their four
      states, s.t. $p_A[2i+x] = p^{i,x}_A$, $q_B[2j+y] = q^{j,y}_B$
    # $p_{\rm dark}$ is the dark count probability per detector
    # $\eta$ is the overall transmissivity
    # $l$ is the Alice-Charlie distance (same for Bob)
    #
    # Probability of losing a photon
    $p_0 = 1 - \eta 10^{-0.2l/20}$
    # Extract protocol statistics
    $\vec{p}_{\rm det}, \hat{\gamma}$= stats($\rho_A, \sigma_B, p_A, q_B, p_0, p_{\rm dark}$)
    # Key generation detection probability
    $p^{0,0}_{\rm det} = \sum_{i=0}^{3} \vec{p}_{\rm det}[i]$
    # Bit error rate
    $e_Z = \vec{p}_{\rm det}[1] + \vec{p}_{\rm det}[2]$
    # Solving for Eve's Gramian matrix [Eq. (3) of main text]
    $\vec{e} = \hat{\gamma}^{-1}\vec{p}_{\rm det}$
    # Phase error rates
    $e_- = {\rm emin}(\rho_A, \sigma_B, p_A, q_B, \vec{e}, e_Z, p^{0,0}_{\rm det})$
    $e_+ = {\rm eplus}(\rho_A, \sigma_B, p_A, q_B, \vec{e}, e_Z, p^{0,0}_{\rm det})$
    # Key rate
    $R = p^{0,0}_{\rm det}\big[1 - h_2(e_Z) - e_Z h_2\big(\frac{1+e_-/e_Z}{2}\big) - (1-ez)h_2\big(\frac{1-(e_++e_Z)/2}{1-e_Z}\big)\big]$
    return R
end function
```

The states $|\xi^\delta\rangle$ parametrized by δ are a model for Alice and Bob attempting to prepare $\{|H\rangle, |V\rangle, {}^{|H\rangle+|V\rangle}\!/\!_{\sqrt{2}}, {}^{|H\rangle-i|V\rangle}\!/\!_{\sqrt{2}}\}$, but each state is subject to a different, constant state-dependent modulation error. The pure $|\xi^\delta\rangle$ states and the resulting key rates were considered in the loss-tolerant protocol [8]. Additionally, were the modulation error a random variable subject to a distribution on the Bloch sphere, we

Algorithm 2: Protocol statistics function

```
function STATS($\rho_A, \sigma_B, p_A, q_B, p_0, p_{\rm dark}$)
    # Loop over all 16 combinations of states in lists: $i$, $j$=0,1; $x$,
      $y$=0,1
    # Probability of passing if both photons arrive
    $p_{\rm pass}[8i+4j+2x+y] = p_A[2i+x]q_B[2j+y]Tr(\rho_A[i,x]$
      $\sigma_B[j,y]\,|\Phi^+\rangle\,\langle\Phi^+|_{AB})$
    # Detection probability including dark counts and loss
      $p_{\rm det}[8i+4j+2x+y] = (1-p_0)^2(1-p_{\rm dark})^2 p_{\rm pass}$
    $[8i+4j+2x+y]$
    $p_{\rm det}[8i+4j+2x+y]+ =2p_A[2i+x]q_B[2j+y]$
      $[p_0^2 p_{\rm dark}^2(1-p_{\rm dark})^2 + p_0(1-p_0)p_{\rm dark}(1-p_{\rm dark})^2]$
    # Filling the 16 rows of the $\hat{\gamma}$ matrix
    $\hat{\gamma}[8i+4j+2x+y] = p_A[2i+x]q_B[2j+y]{\rm vec}(\rho_A[i,x]\sigma_B[j,y])$
    return $\vec{p}_{\rm det}, \hat{\gamma}$
end function
```

Algorithm 3: Phase errors

```
**function** EMIN($\rho_A, \sigma_B, p_A, q_B, \vec{e}, e_Z, p_{\rm det}^{0,0}$)
    # Reshape $\vec{e}$ into a matrix
    $\hat{e} = $ reshape($\vec{e}$)
    # We use the CVXPY and MOSEK packages for solving semidefinite programs
    import CVXPY
    # Define the 8 × 8 Gramian matrix from Eq. (C11) for the $A'B'$ systems as the optimization variables of the system
    $G = $ CVXPY.Variable((8,8))
    # Define a list of constraints on G, such as PSD and constraint from Eq. (C12)
    constraints $= [G \succeq 0]$
    # For $x = 0, 1; m, m', n, n' = 0, 1$
    constraints += $[G[4x + 2m + n, 4x + 2m' + n'] = p_A[x]q_B[(x + 1) \bmod 2]\rho_A[0, x][m, m']\sigma_B[0, (x + 1) \bmod 2][n, n']]$
    # Define the objective function
    $e_- = -\frac{2}{p_{\rm det}^{0,0}} \sum_{m,m',n,n'} {\rm Re}(\hat{e}[2m + n, 2m' + n']G[2m + n, 4 + 2m' + n'])$
    constraints += $[e_- \geqslant 0, e_- \leqslant e_Z]$
    # Use CVXPY to solve problem
    prob = CVXPY. Problem (CVXPY.Maximize($e_-$),constraints)
    prob.solve(solver = CVXPY. MOSEK)
    $e_- = $ prob.value
    **return** $e_-$
**end function**

**function** EPLUS($\rho_A, \sigma_B, p_A, q_B, \vec{e}, e_Z, p_{\rm det}^{0,0}$)
    # Reshape $\vec{e}$ into a matrix
    $\hat{e} = $ reshape($\vec{e}$)
    # We use the CVXPY and MOSEK packages for solving semidefinite programs
    import CVXPY
    # Define the 8 × 8 Gramian matrix from Eq. (C13) for the $A'B'$ systems as the Variable of the system
    $G = $ CVXPY.Variable((8,8))
    # Define a list of constraints on G, such as PSD and constraint from Eq. (C14)
    constraints $= [G \succeq 0]$
    #For $x = 0, 1; m, m', n, n' = 0, 1$
    constraints += $[G[4x + 2m + n, 4x + 2m' + n'] = p_A[x]q_B[x]\rho_A[0, x][m, m']\sigma_B[0, x][n, n']]$
    # Define the objective function
    $e_+ = 1 - \frac{2}{p_{\rm det}^{0,0}} \sum_{m,m',n,n'} {\rm Re}(\hat{e}[2m + n, 2m' + n']G[2m + n, 4 + 2m' + n'])$
    constraints += $[e_+ \geqslant e_Z, e_+ \leqslant 1]$
    # Use CVXPY to solve problem
    prob = CVXPY. Problem (CVXPY. Minimize($e_+$),constraints)
    prob.solve(solver = CVXPY. MOSEK)
    $e_+ = $ prob.value
    **return** $e_+$
**end function**
```

expect the average state to be mixed with a shorter than unit Bloch vector. This effect is accounted for with the depolarizing channel parametrized by p, which indicates with some probability the maximally mixed state is sent, shortening the Bloch vector. The depolarizing channel can also be used to model any thermal photons that are accidentally produced during state preparation.

[1] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[2] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, Rev. Mod. Phys. **92**, 025002 (2020).

[3] H.-K. Lo, M. Curty, and B. Qi, Phys. Rev. Lett. **108**, 130503 (2012).

[4] S. L. Braunstein and S. Pirandola, Phys. Rev. Lett. **108**, 130502 (2012).

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

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

[7] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, Quantum Inf. Comput. **4**, 136 (2003).

[8] K. Tamaki, M. Curty, G. Kato, H.-K. Lo, and K. Azuma, Phys. Rev. A **90**, 052314 (2014).

[9] M. Pereira, M. Curty, and K. Tamaki, npj Quantum Inf. **5**, 62 (2019).

[10] W. Wang, K. Tamaki, and M. Curty, arXiv:2001.8086.

[11] Á. Navarrete, M. Pereira, M. Curty, and K. Tamaki, arXiv:2007.3364.

[12] Z.-Q. Yin, C.-H. F. Fung, X. Ma, C.-M. Zhang, H.-W. Li, W. Chen, S. Wang, G.-C. Guo, and Z.-F. Han, Phys. Rev. A **88**, 062322 (2013).

[13] I. Devetak and A. Winter, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **461**, 207 (2005).

[14] P. J. Coles, E. M. Metodiev, and N. Lütkenhaus, Nat. Commun. **7**, 11183 (2016).

[15] A. Winick, N. Lütkenhaus, and P. J. Coles, Quantum **2**, 77 (2018).

[16] H.-K. Lo, Quantum Info. Comput. **1**, 81 (2001).

[17] Z. Tang, K. Wei, O. Bedroya, L. Qian, and H.-K. Lo, Phys. Rev. A **93**, 042308 (2016).

[18] F. Xu, K. Wei, S. Sajeed, S. Kaiser, S. Sun, Z. Tang, L. Qian, V. Makarov, and H.-K. Lo, Phys. Rev. A **92**, 032305 (2015).

[19] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim, Phys. Rev. Lett. **100**, 110502 (2008).

[20] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, IEEE Trans. Inf. Theory **55**, 1898 (2009).

[21] J. M. Renes and G. Smith, Phys. Rev. Lett. **98**, 020502 (2007).

[22] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim, IEEE Trans. Inf. Theory **54**, 2604 (2008).

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

[24] N. T. Islam, C. C. W. Lim, C. Cahall, J. Kim, and D. J. Gauthier, Phys. Rev. A **97**, 042347 (2018).

[25] Y. Wang, I. W. Primaatmaja, E. Lavie, A. Varvitsiotis, and C. C. W. Lim, npj Quantum Inf. **5**, 39 (2019).

[26] I. W. Primaatmaja, E. Lavie, K. T. Goh, C. Wang, and C. C. W. Lim, Phys. Rev. A **99**, 062332 (2019).

[27] N. T. Islam, C. C. W. Lim, C. Cahall, B. Qi, J. Kim, and D. J. Gauthier, Quantum Sci. Technol. **4**, 035008 (2019).

[28] L. Liu, Y. Wang, E. Lavie, C. Wang, A. Ricou, F. Z. Guo, and C. C. W. Lim, Phys. Rev. Appl. **12**, 024048 (2019).

[29] E. Y. Z. Tan, R. Schwonnek, K. T. Goh, I. W. Primaatmaja, and C. C. W. Lim, arXiv:1908.11372.

[30] R. Renner, Security of quantum key distribution (2005), arXiv:quant-ph/0512258.

[31] A. Sit, F. Bouchard, R. Fickler, J. Gagnon-Bischoff, H. Larocque, K. Heshami, D. Elser, C. Peuntinger, K. Günthner, B. Heim, C. Marquardt, G. Leuchs, R. W. Boyd, and E. Karimi, Optica **4**, 1006 (2017).

[32] L. Wang, S.-M. Zhao, L.-Y. Gong, and W.-W. Cheng, Chin. Phys. B **24**, 120307 (2015).

[33] X.-Y. Wang, S.-H. Zhao, C. Dong, Z.-D. Zhu, and W.-Y. Gu, Quant. Info. Proc. **18**, 304 (2019).

[34] Y.-L. Tang, H.-L. Yin, S.-J. Chen, Y. Liu, W.-J. Zhang, X. Jiang, L. Zhang, J. Wang, L.-X. You, J.-Y. Guan, D.-X. Yang, Z. Wang, H. Liang, Z. Zhang, N. Zhou, X. Ma, T.-Y. Chen, Q. Zhang, and J.-W. Pan, Phys. Rev. Lett. **113**, 190501 (2014).

[35] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. (Cambridge University Press, New York, 2011).

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

[37] S. Diamond and S. Boyd, Journal of Machine Learning Research **17**, 1 (2016).

[38] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, Journal of Control and Decision **5**, 42 (2018).

[39] M. Grant and S. Boyd, " CVX: Matlab software for disciplined convex programming, Version 2.1," http://cvxr.com/cvx (2014).

[40] M. Mafu, A. Dudley, S. Goyal, D. Giovannini, M. McLaren, M. J. Padgett, T. Konrad, F. Petruccione, N. Lütkenhaus, and A. Forbes, Phys. Rev. A **88**, 032305 (2013).

[41] D. Cozzolino, D. Bacco, B. Da Lio, K. Ingerslev, Y. Ding, K. Dalgaard, P. Kristensen, M. Galili, K. Rottwitt, S. Ramachandran, and L. K. Oxenløwe, Phys. Rev. Appl. **11**, 064058 (2019).

[42] L. Sheridan and V. Scarani, Phys. Rev. A **82**, 030301(R) (2010).