

Measurement-device-independent quantum key distribution with passive time-dependent source side channels

J. Eli Bourassa^{1,*}, Amita Gnanapandithan^{2,*}, Li Qian^{2,3} and Hoi-Kwong Lo^{1,2,3,4}

¹*Department of Physics, University of Toronto, Toronto, Canada*

²*Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada*

³*Center for Quantum Information and Quantum Control, University of Toronto, Toronto, Canada*

⁴*Department of Physics, University of Hong Kong, Hong Kong*



(Received 16 October 2022; accepted 7 December 2022; published 20 December 2022)

While measurement-device-independent (MDI) quantum key distribution (QKD) allows two trusted parties to establish a shared secret key from a distance without needing to trust a central detection node, their quantum sources must be well characterized, with side channels at the source posing the greatest loophole to the protocol's security. In this paper, we identify a time-dependent side channel in a common polarization-based QKD source that employs a Faraday mirror for phase stabilization. We apply a recently developed numerical proof technique [Phys. Rev. A **99**, 062332 (2019)] to quantify the sensitivity of the secret key rate to the quantum optical model for the side channel, and to develop strategies to mitigate the information leakage. In particular, we find that the MDI three-state and BB84 protocols, while yielding the same key rate under ideal conditions, have diverging results in the presence of a side channel, with BB84 proving more advantageous. While we consider only a representative case example, we expect the strategies developed and key rate analysis method to be broadly applicable to other leaky sources.

DOI: [10.1103/PhysRevA.106.062618](https://doi.org/10.1103/PhysRevA.106.062618)

I. INTRODUCTION

Quantum key distribution (QKD) allows two separated parties to have information-theoretic secure communication by leveraging the principles of quantum mechanics [1]. In practice, QKD systems suffer from imperfections and open up side-channel attacks. In particular, detectors are the weakest link in QKD. Luckily, measurement-device-independent (MDI) QKD removes all side channels in detectors [2]. Nonetheless, imperfections of the quantum state source continue to threaten the security of MDI protocols. To partly address this challenge, the loss tolerant protocol [3] provides a proof technique for dealing with state preparation flaws, with extensions of the proof available to account for the decoy state method [4], and mixed states [5]. The method from [6] can also treat flawed sources, under the condition the encoded signals remain confined to a qubit space. Unfortunately, these methods to deal with state preparation flaws only account for systematic errors in the two-dimensional (qubit) degree of freedom that Alice and Bob intentionally encode, meaning these techniques are not sufficient to account for source side channels. In an effort to generalize to sources that do not output idealized qubits, recent security proof techniques have been developed to deal with sources leaking decoy state parameters [7,8], and encoding information, with analytic approaches given in [9,10] and numerical techniques in [11,12].

With such security proof techniques now available, it is time they be applied to develop practical strategies for MDI QKD protocols employing realistic sources, bringing closer together the gap between idealized security proofs and experimental realities. In this paper, we study a common optical source for polarization-based MDI QKD which relies on a Faraday mirror for polarization stabilization [4,13–18]. We determine that this experimental setup introduces a passive side channel (i.e., not introduced by Eve) due to leakage light between optical pulses being unintentionally modulated in a time-dependent manner, a loophole that has not been identified in the literature to the best of our knowledge. While some work has been done on computing secure key rates in the presence of passive side channels—the authors of [19] establish a lower bound on the key rate in the presence of passive side channels using signal indistinguishability—most previous works focus on active side channels (side channels introduced by Eve), such as Trojan horse attacks [20–22], despite passive side channels being harder to avoid. Note that the passive-active dichotomy here refers to whether Eve introduces the side channel, as opposed to the passive-active distinction made for optical elements [23]. For an exhaustive review of hacking strategies for various QKD systems, including via side channels, see Table I of [24].

In this particular passive side-channel case, we are faced with the seemingly daunting task of incorporating optical states distributed over a continuum of temporal modes into a security proof. However, we find that the versatile proof technique from [12] can be employed even in this scenario, a modest extension of its already wide applicability. As a numerical approach, the technique from [12] is particularly

*These authors contributed equally to this work.

†eli.bourassa@mail.utoronto.ca

‡amita.gnanapandithan@mail.utoronto.ca

well suited to our task, as it allows one to integrate detailed information about the initial states sent by Alice and Bob (including the time-dependent side channels) and all the observed detection statistics in the protocol as constraints in the security proof.

Using the Faraday mirror source as a representative case example of sources with time-varying side channels, we calculate the secret key rate under various assumptions and scenarios to better determine strategies for mitigating the information leaked via the side channel. In particular, we investigate how the model for the state of the side channel can have a significant impact on the amount of key generated, reinforcing the importance of carefully characterizing the optical output of the source. We present a few practical strategies for increasing the key rate, such as using all available detection statistics, sending more states than what would be required in the ideal protocol, and optimizing the choice of which test state to send from the Bloch sphere. As part of this analysis, we determine that while the MDI three-state protocol [3,25] yields the same key rates as the MDI BB84 protocol [2,26] in the ideal case of no side channel, in the presence of leakage light these two protocols diverge, with BB84 being the more advantageous choice.

We briefly summarize why we will employ the numerical approach of [12] based on semi-definite programming to address the side-channel problem. Note that some other approaches such as the loss-tolerant protocol approach [3] and uncharacterized qubit approach [6] cannot be applied to the side-channel problem because those approaches assume the optical source sends out a qubit and such a qubit assumption is violated by side channels.

On the other hand, approaches such as [11] and the reference state approach [9,10] do work for side channels and nonqubit sources. Nonetheless, we find that technique from [11] relaxes the task of bounding the phase error to a linear program, and, therefore, it gives a less strict result than using the approach in [12]. As for the reference state technique [9,10], we find that it gives a worse key rate for the MDI version of BB84 protocol in the presence of side channels than the approach in [12]. For these reasons, we find that the approach in [12] is highly suitable for addressing the side-channel problem.

The structure of this paper is as follows: in Sec. II we review the security proof technique from [12] and compare it to competing proof techniques [3,6,9–11] to justify our choice of approach. Then in Sec. III we study the case example of a polarization-based MDI QKD setup that employs a Faraday mirror in the transmitter for polarization stabilization; here we identify a side channel arising from the time-varying polarization modulation of the leakage light between optical pulses. This polarization modulation is correlated with the polarization encoding of Alice and Bob's signals. We determine that the proof technique from [12] can be applied to treat practical sources with time-varying optical signals.

Finally, in Sec. IV we provide key rate results for various protocol scenarios. We determine the impact of the source model, finding that accounting for the time-dependent nature of the side channel provides a benefit to the key rate over more pessimistic, rudimentary models, thus demonstrating the importance of careful side-channel characterization.

Additionally, we find a divergence between the three-state and BB84 protocols in the presence of a side channel, and determine that the extra state sent in BB84, while redundant under ideal conditions, is able to better mitigate the source information leakage.

II. BACKGROUND

To understand the dependence of the secret key rate on the side channel, we first provide some background on the components that are required for the key rate calculation. In Sec. II A we review the proof technique from [12], and in Sec. II B, we compare our choice of proof technique to other potential options we could have chosen. As our key rate calculations rely on the decoy state protocol, we provide a review of its use in MDI QKD in Appendix A.

A. Security proof technique based on semidefinite programming

Semidefinite programs (SDPs) are a class of convex optimization problems that can be written in the form

$$\begin{aligned} \text{maximize} \quad & f_0(G) = \text{Tr}(A_0 G) \\ \text{s. t.} \quad & f_i(G) = \text{Tr}(A_i G) \geq b_i, \quad i = 1, \dots, m \\ & G \geq 0, \end{aligned}$$

where G is a positive semidefinite (PSD) matrix (i.e., has nonnegative eigenvalues) whose elements form the optimization variables of the problem. $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function we seek to maximize. $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are the constraint functions, and b_i are the constraint bounds. Importantly, the objective and the constraint functions are all linear functions of the elements of G , with the coefficients of the linear functions contained in the matrices A_i . SDPs are increasingly being used for QKD security proofs [12,22,27–33], due in part to the availability of fast and mature numerical implementations of solvers [34,35].

In [12] the authors present a versatile numerical proof technique based on semidefinite programs (SDPs) for MDI QKD protocols. The objective function of the SDP is the phase error of the key rate formula—either Shor-Preiskill [36] or the Gottesman, Lo, Lutkenhaus, and Preiskill (GLLP) version for decoy states [37] [see Eq. (A6)]—meaning an optimal solution provides a secure lower bound. The constraints are provided by the detection statistics of the protocol, as well as initial state information, which is especially useful since it allows more experimental information to be used to quantify security. Here we review the proof technique from [12] so that later we can apply it to the case of an MDI QKD protocol with a source side channel.

To begin, we distinguish between the full optical states that Alice and Bob send to Charlie from the components of those states from which we will derive security. We will refer to the components of the optical states used to derive security as the *signal* states. As an example, in an ideal decoy state protocol, the full optical states are phase-randomized weak coherent pulses (WCPs) with different intensities, while the signal states used to derive security are the single-photon components. We review the decoy state protocol in Appendix A. Alternatively, in an ideal phase-encoding protocol that does

not use decoy states, the optical state and the signal state are one and the same.

Let the signal states that Alice prepares be denoted by $|\psi_x^i\rangle_A$, where (i, x) indicate her choice of basis and bit value. Analogously, we can write Bob's states as $|\varphi_y^j\rangle_B$, with his basis and bit choice given by (j, y) . Like in [12], we will assume that the signal states are pure states; however, a path to treating mixed states is available via the technique from [5]. Note that, even though phase-randomized WCPs in an ideal decoy state protocol are mixed states, the signal states (single-photon components) are pure states. At a high level, in an MDI QKD protocol, Alice and Bob send their signal states to Eve, who in turn makes an announcement z conditioned on a measurement she may or may not execute faithfully. For the sake of simplicity, we assume $z = P, F$, corresponding to a binary pass or fail outcome, but this can be generalized to account for more announcements.

Since quantum mechanics obeys unitary evolution, we can write the evolution of the joint state as

$$|\psi_x^i \varphi_y^j\rangle_{AB} \rightarrow \sum_{z=P,F} |e_{x,y,z}^{i,j}\rangle_E |z\rangle_Z, \quad (1)$$

where $|e_{x,y,z}^{i,j}\rangle_E$ are subnormalized vectors that can be used to completely characterize Eve's state. Were we to know the states $|e_{x,y,z}^{i,j}\rangle_E$, we would have full knowledge of Eve's information about the key, and in turn could calculate the key rate exactly. Alas, the exact state of Eve's system is generally unknown; however, we can impose constraints on the state vectors $|e_{x,y,z}^{i,j}\rangle_E$. As we will review now, one can frame the secret key rate calculation as a semidefinite program, where the PSD matrix used as an optimization variable is the Gram matrix of Eve's vectors which we denote G_E . We recall that the elements of a Gram matrix for a set of vectors are all the pairwise inner products of the vectors, meaning G_E has elements $\langle e_{x',y',z'}^{i',j'} | e_{x,y,z}^{i,j} \rangle_E$. Gram matrices are always PSD.

The first type of constraint comes from the unitary evolution of states in quantum mechanics; namely, the inner product structure of the initial states must be preserved in the final states [12]. If Alice and Bob each have n_A and n_B basis choice settings, each basis choice associated with two bit choices, then the inner product constraint yields $(n_A \times n_B \times 2 \times 2)^2$ constraints of the form

$$\langle \psi_{x'}^{i'} \varphi_{y'}^{j'} | \psi_x^i \varphi_y^j \rangle_{AB} = \sum_z \langle e_{x',y',z}^{i',j'} | e_{x,y,z}^{i,j} \rangle_E, \quad (2)$$

where the fact that the announcements are classical means $\langle z | z' \rangle_Z = \delta_{z,z'}$. Note that these constraints are linear in the elements of G_E .

The next type of constraint on G_E comes from the observed detection statistics [12]. Let the probability of Eve announcing a successful detection event, conditioned on Alice and Bob having chosen basis and bit choices (i, j, x, y) be denoted by $p_{\text{pass}}^{i,j,x,y}$. In an ideal decoy state protocol, $p_{\text{pass}}^{i,j,x,y} \equiv p_{\text{pass},1,1}^{i,j,x,y}$, since the single-photon components are the signal states. This can be related to the elements of G_E as follows:

$$p_{\text{pass}}^{i,j,x,y} = \langle e_{x,y,P}^{i,j} | e_{x,y,P}^{i,j} \rangle_E. \quad (3)$$

We recall the states are subnormalized so we do not expect the above equation to be equal to 1. If the signal states are the

same as the full optical states, then $p_{\text{pass}}^{i,j,x,y}$ would be directly observable in practice. Alternatively, if one is performing a decoy state MDI QKD protocol, then, as reviewed in Appendix A, one first uses the statistics of the full optical states in a linear program to establish upper and lower bounds on $p_{\text{pass}}^{i,j,x,y}$:

$$p_{\text{pass},L}^{i,j,x,y} \leq p_{\text{pass}}^{i,j,x,y} \leq p_{\text{pass},U}^{i,j,x,y}. \quad (4)$$

Thus, depending on the protocol, one either obtains $4n_A n_B$ equality constraints, or $8n_A n_B$ inequality constraints on G_E . Like the previous set of constraints, these are also linear in elements of G_E .

So far, we have identified the Gram for Eve's system G_E as a PSD matrix, as well as various linear constraints on its elements. We now review how to write the phase error as the objective function of an SDP. To start, we will assume that the basis choice $(i, j) = (0, 0)$ corresponds to the key generation basis. Moving to a virtual picture, we can think of Alice and Bob's signal states being entangled with virtual qubits \bar{A} and \bar{B} that they keep in their laboratory:

$$|\Psi_{\text{virt}}\rangle_{\bar{A}\bar{B}AB} = \sum_{x,y=0}^1 |x\rangle_{\bar{A}} |y\rangle_{\bar{B}} |\psi_x^0 \varphi_y^0\rangle_{AB}, \quad (5)$$

where measurement of $\bar{A}\bar{B}$ in the computational basis yields the bit values of the secret key. Let the virtual state evolve to $|\Psi_{\text{virt}}\rangle_{\bar{A}\bar{B}AB} \rightarrow |\Gamma\rangle_{\bar{A}\bar{B}EZ}$ with

$$|\Gamma\rangle_{\bar{A}\bar{B}EZ} = \sum_{x,y=0}^1 |x, y\rangle_{\bar{A}\bar{B}} \sum_{z=P,F} |e_{x,y,z}^{0,0}\rangle_E |z\rangle_Z, \quad (6)$$

since we used Eq. (1).

Through the process of sending their key generation signal states to Eve (who conducts a measurement), as well as postselecting on $z = P$, Alice and Bob end up with a mixture of Bell states in the $\bar{A}\bar{B}$ virtual qubits. The virtual picture therefore allows us to formally define the phase error rate that characterizes security in the key rate in Eq. (A6). Assuming, without loss of generality, that the target Bell state of the protocol is $|\Phi^+\rangle_{\bar{A}\bar{B}} = \frac{1}{\sqrt{2}}(|00\rangle_{\bar{A}\bar{B}} + |11\rangle_{\bar{A}\bar{B}})$, then the phase error rate is defined to be the probability that the $\bar{A}\bar{B}$ virtual qubits held by Alice and Bob end up in Bell states with the incorrect phase (see derivation in Appendix B):

$$e_{ph} = \frac{\langle \Gamma | (\mathbb{M}_{\bar{A}\bar{B}}^- \otimes |P\rangle\langle P|_Z) | \Gamma \rangle_{\bar{A}\bar{B}EZ}}{\langle \Gamma | (|P\rangle\langle P|_Z) | \Gamma \rangle_{\bar{A}\bar{B}EZ}} = \frac{1}{2} - \frac{\text{Re}(\langle e_{0,0,P}^{0,0} | e_{1,1,P}^{0,0} \rangle_E + \langle e_{0,1,P}^{0,0} | e_{1,0,P}^{0,0} \rangle_E)}{\sum_{x,y} p_{\text{pass}}^{0,0,x,y}}, \quad (7)$$

with

$$\mathbb{M}_{\bar{A}\bar{B}}^- = (|\Phi^-\rangle\langle\Phi^-| + |\Psi^-\rangle\langle\Psi^-|)_{\bar{A}\bar{B}} \quad (8)$$

representing the measurement operator associated with Bell states of the wrong phase to the target $|\Phi^+\rangle_{\bar{A}\bar{B}}$. Note that e_{ph} is a linear function of the elements of G_E as required for an SDP. With the additional constraint that $0 \leq e_{ph} \leq 1/2$, such that we are within the region where the binary entropy function increases monotonically, then we can maximize e_{ph} via an SDP and determine a secure lower bound on the key

rate using the Shor-Preskill formula [36] or the key rate from (A6). Note that in a decoy state protocol without leakage light, $e_{ph} = e_{ph,1,1}$, since the signal states correspond to the single-photon components.

In summary, [12] provides a method for obtaining a secure lower bound on the key rate using an SDP of the form

$$\begin{aligned} & \text{maximize} && e_{ph} \\ & \text{s.t.} && \langle \psi_{x'}^i \varphi_{y'}^{j'} | \psi_x^i \varphi_y^j \rangle_{AB} = \sum_z \langle e_{x',y',z}^{i',j'} | e_{x,y,z}^{i,j} \rangle_E \\ & && P_{\text{pass},L}^{i,j,x,y} \leq \langle e_{x,y,P}^{i,j} | e_{x,y,P}^{i,j} \rangle_E \leq P_{\text{pass},U}^{i,j,x,y} \\ & && 0 \leq e_{ph} \leq 1/2 \\ & && G_E \geq 0. \end{aligned}$$

With this technique in hand, we apply it to the case example of an MDI QKD source with a side channel, which we will describe in the next section. First, however, we justify our choice of proof technique by comparing it with competing methods.

B. Comparison to competing proof techniques

Given our review of the numerical approach from [12], it is worth comparing with competing proof techniques for MDI QKD to see why the approach we have chosen is well-suited to the problem we wish to study. The loss-tolerant protocol [3] and the proof technique from [6] are both leading techniques for quantifying security in the presence of state-preparation flaws, the former requiring knowledge of the initial states, while the latter can simply use the detection statistics. However, in both techniques, one needs to assume that the optical source is outputting a qubit state, meaning they are insufficient to treat scenarios involving source side channels, since the extra state sent with the encoded qubit, e.g., an optical coherent state, can easily break the assumption that the source only outputs states from a two-dimensional Hilbert space.

The numerical proof technique developed in [27,28] also uses SDPs to compute the secret key rate; however, they work directly with the Devetak-Winter key rate formula [38], as opposed to the Shor-Preskill key rate [36]. Working with the Devetak-Winter key rate requires solving a series of SDPs, which is more cumbersome and numerically slower than a direct calculation of the phase error.

Two generalizations of the loss tolerant protocol have been developed to deal with nonqubit sources, in part for the purpose of studying source side channels: the technique from [11] and the reference state technique [9,10]. These techniques are directly comparable to [12] as they all use the Shor-Preskill key rate, with the core task of the proof being to find an upper bound on the phase error rate. While shown to perform more poorly than the reference state technique [9], the technique from [11] may be the most directly comparable to [12]. Both methods allow one to consider an arbitrary number of initial states that are not confined to a qubit space, and involve using all observed detection statistics. As we show in Appendix C, however, the approach from [11] relaxes the task of bounding the phase error to a linear program. The constraints of the linear program are provided by the detection probabilities and the initial states; however, fewer constraints

are provided by the initial states than the approach from [12], since the overlaps between states with different (i, j, x, y) are not considered. Moreover, linear programs are a class of convex optimization problems contained within SDPs, meaning the constraints on the optimization variables used to compute the phase error are less strict than solving the full SDP as done in [12]. Thus, we expect the key rates provided by the SDP numerical approach to be greater or equal to those calculated using the technique from [11] in general. In Appendix C we provide the illustrative example of the three-state protocol with a side channel to explicitly demonstrate the key rates provided by the SDP method outperform those from [11].

Finally, while the reference state technique is a purely analytic approach, a disadvantage is that it can currently only treat the case when Alice and Bob each sends three states. The crux of the technique is to consider hypothetical detection statistics and phase error stemming from a fictitious set of reference states, and then bound the actual phase error of the protocol using the real detection statistics and the deviation between the reference and real states [9,10]. In particular, in [9] the strategy for treating protocols involving four states, such as BB84, is to consider random alternation between two three-state protocols, where each of the X-basis BB84 states act as the third state. While the three-state protocol and BB84 yield the same key rates when the initial states are qubits [3], one of the observations that this work will provide is that information from the seemingly redundant fourth state of BB84 can provide extra constraints to boost the key rate in the presence of a source side channel (i.e., when the qubit assumption is broken). Thus, a downside of using the reference state technique for protocols involving more than three states is that the key rate calculation will only ever be constrained by the statistics and initial states of three out of the four states, which leaves valuable information on the table, at the cost of higher key rate.

III. SOURCE SIDE CHANNELS: A CASE EXAMPLE

Having reviewed the necessary components for the security proof, we now study a case example of an MDI QKD source with a side channel. In Sec. III A we identify a time-dependent passive source side channel which occurs when using a Faraday mirror for stable electro-optic bit modulation [4,13–18]. We provide quantum optical modeling of the side channel in Sec. III B, and in Sec. III C, we link the model to the security proof technique described in Sec. II to assess its impact on security while taking its time-dependent nature into account.

A. Origin of the side channel

Several polarization and phase encoding implementations of MDI, prepare-and-measure, and plug-and-play QKD make use of an electro-optic phase modulator and Faraday mirror for optical bit modulation [4,13,39]. The Faraday mirror is added, as shown in Fig. 1, to remove the temperature dependence of the phase modulator [4,13–17]. Optical pulses first travel forward through the PM, copropagating with a voltage pulse. During this first trip, they experience both voltage and unintentional temperature induced phase modulation. After reflection from the Faraday mirror, the pulses travel back

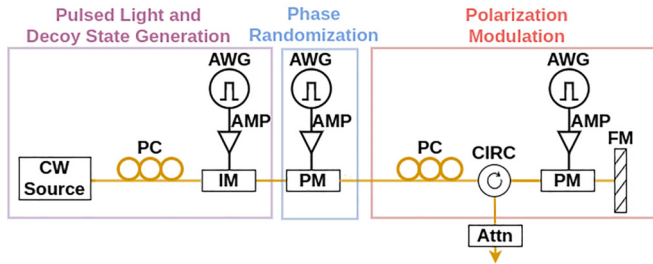


FIG. 1. Experimental setup for polarization encoding MDI-QKD transmitter. An intensity modulator is used to create pulsed light (including decoy states) from a continuous wave light source. Then the pulses go through a phase randomization unit, followed by a polarization modulation (encoding) unit. PC, polarization controller; IM, intensity modulator; AMP, voltage amplifier; AWG, arbitrary waveform generator; PM, phase modulator; CIRC, optical circulator; FM, Faraday mirror; Attn, optical attenuator.

through the PM, such that they do not collide with any counterpropagating voltage pulses. Hence, during this second trip, they experience only temperature-induced phase modulation. Although the usage of a Faraday mirror drastically reduces state preparation flaws, we found that it creates a source side channel.

This side channel occurs due to the presence of weak light leakage between the optical pulses into which bits are encoded. These optical pulses are carved out from continuous wave light using an electro-optic intensity modulator (IM). Due to the finite extinction ratio of pulses that can be created with an IM, the presence of weak light leakage is inevitable. Of course, the phase of this leakage light is not intentionally modulated. In other words, no voltage is applied to the phase modulator as this light travels through it the first time. However, after reflection from the Faraday mirror, some of this leakage light would inevitably collide with a counterpropagating voltage pulse within the phase modulator. Therefore, this leakage light would experience unintentional voltage-induced phase modulation, forming a source side channel whose impact on security must be quantified.

We find the modulation of the leakage light to be time-dependent, as it travels in the opposite direction of the voltage pulse. It is also dependent on the voltage pulse shape used for phase modulation and the phase modulator electrode length. For our particular setup, the time dependence is shown in Fig. 2. Refer to Appendix D for further details on how Fig. 2 was derived.

Several simplifying assumptions regarding the optical and leakage signals are made when performing the security analysis.

(1) In our particular setup (see Fig. 1), polarization encoding occurs after setting the decoy state intensity and performing phase randomization. Hence, we assume that the decoy state intensity and phase randomization of optical pulses (encoded signals) are independent of their polarization encoding.

(2) We assume that the voltage pulses delivered to the polarization modulation PM are square pulses, such that the polarization of encoded pulses are time-independent.

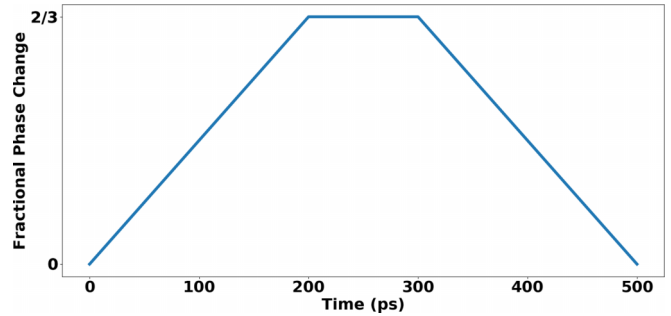


FIG. 2. Time-dependent fractional phase change applied to the leakage light. Fraction is with respect to the phase change applied to the corresponding encoded light.

(3) In our experimental setup, there is a minimal correlation between the decoy state intensity and leakage light intensity. This correlation stems from the tails of the pulse shaping voltage pulses, which are small compared to the full duration of leakage light between pulses. Therefore, we assume that leakage signals carry no information about the decoy state intensity setting of encoded signals.

(4) We assume that the voltage pulses strictly overlap with the optical pulses within the phase randomizing PM. Hence, the leakage signals carry no information about the phase randomization of encoded signals.

(5) We assume that the leakage intensity is uniform in time, but the method could also easily treat time-varying intensity.

Note that these assumptions could be broken and could be incorporated into the security analysis using techniques from [8]. However, in our particular experimental setup, they are not the leading-order source of information leakage, which we take to be the unintentional polarization modulation of the leakage light after reflection in the Faraday mirror. We can now proceed to model the quantum state of the source's output light.

B. Quantum optical modeling

To proceed with the security proof technique from Sec. II A, we must first model the states transmitted by the source so that we can compute the inner products of the signal states. The full optical state can be written as a separable state of Alice's and Bob's transmitted states: $\rho_A^{k,i,x} \otimes \sigma_B^{l,j,y}$. The parameters (k, l) refer to their decoy state intensity settings, (i, j) refer to their basis choice, and (x, y) refer to their bit choices. The basis and bit choices impact the polarization set with the polarization phase modulator and are independent from the choice of intensity setting.

Alice's state can be further broken down into the side-channel state which represents the leakage light, and the encoding state which represents the optical pulses into which the basis and bit information are intentionally encoded:

$$\rho_A^{k,i,x} = \rho_{\text{enc}}^{k,i,x} \otimes \rho_{\text{leak}}^{i,x}. \quad (9)$$

We assume that the intensity modulation and phase randomization are timed with the optical pulses, and that in between pulses when the leakage light is passing through

these modules, no phase randomization is applied, and the intensity modulation attenuates the light as much as possible to constant minimum but nonzero intensity. This has several consequences: first, this means that only $\rho_{\text{enc}}^{k,i,x}$ carries information about the intensity setting, and that $\rho_{\text{leak}}^{i,x}$ carries information about neither the random phase nor the intensity. This means we can treat $\rho_{\text{enc}}^{k,i,x}$ as a perfectly phase-randomized WCP just as in an ideal MDI decoy state protocol without leakage light. Second, our assumptions mean that we can treat $\rho_{\text{leak}}^{i,x}$ as a pure state:

$$\rho_{\text{leak}}^{i,x} = |\chi_x^i\rangle\langle\chi_x^i|_{\text{leak}}. \quad (10)$$

We now move to model the polarization module of the source and the time-dependent nature of the side-channel state. When $\rho_{\text{enc}}^{k,i,x}$ passes through the phase modulator, the controlling voltage pulse is timed with the optical signal such that the resulting polarization, determined by settings (i, x) , is time-independent across the length of the optical pulse. By contrast, because $\rho_{\text{leak}}^{i,x}$ is traveling in the opposite direction, it acquires a time-dependent polarization. Let the creation operator for a photon at time t with polarization angles $\theta_x^i(t)$ and $\phi_x^i(t)$ be given by

$$a_{t,i,x}^\dagger = \cos[\theta_x^i(t)]a_{t,H}^\dagger + \sin[\theta_x^i(t)]e^{i\phi_x^i(t)}a_{t,V}^\dagger, \quad (11)$$

where H and V denote the horizontal and vertical polarization modes, with the raising and lowering operators satisfying $[a_{t,m}, a_{t,m'}^\dagger] = \delta(t-t')\delta_{m,m'}$, $m = H, V$. As non-phase-randomized laser light, the leakage light at a given instant in time t can be treated as a coherent state with amplitude $\alpha_x^i(t)$, meaning over multiple times, the state can be written in general as

$$|\alpha_x^i\rangle = \exp\left\{\int dt[\alpha_x^i(t)a_{t,i,x}^\dagger - \alpha_x^{i*}(t)a_{t,i,x}]\right\}|\text{vac}\rangle. \quad (12)$$

The above equation can be understood as a tensor product of kets defining the state of each temporal mode associated with $a_{t,i,x}^\dagger$. The state defined by each ket is a coherent state, and by varying the amplitude $\alpha_x^i(t)$, the resulting state is a coherent state with varying intensity and polarization defined over a continuum of temporal modes, i.e., a time-dependent state. In the source we are studying, we assume the leakage light has an effectively constant intensity $|\alpha_0|^2$ and polar angle θ , while the azimuthal angle ϕ changes with time based on the interaction with the phase modulator, as shown in Fig. 2; however, the technique we will apply could easily be used to study time-dependent intensity and polar angles as this would just modify the integral over time used to calculate the inner product between two side-channel states. The state of the leakage light associated with a given pulse is spread over multiple temporal modes, and is given by

$$|\chi_x^i\rangle_{\text{leak}} = \exp\left[\int_{-\Delta/2}^{\Delta/2} dt \alpha_0 a_{t,i,x}^\dagger - \alpha_0^* a_{t,i,x}\right]|\text{vac}\rangle, \quad (13)$$

where $a_{t,i,x}^\dagger$ here denotes the creation operator for a polarized photon with time-independent polar angle θ_x^i and time-dependent azimuthal angle $\phi_x^i(t)$ as in Fig. 2, where the angles depend on Alice's basis and bit choices (i, x) . Δ

is the duration of the leakage light that contains encoding information, which from Fig. 2 is 500 ps.

In summary, we have that the output state of Alice's source can be treated as a tensor product of a perfectly phase-randomized WCP in the encoded mode with a time-varying pure coherent state in the side-channel mode. The time-varying polarization of the side-channel state depends on the basis and bit values chosen, but not the intensity setting or random phase used for the decoy state method. We can model Bob's source in the same way, denoting his encoded and leakage states by $\sigma_{\text{enc}}^{l,j,y}$ and $|\zeta_y^j\rangle_{\text{leak}}$, respectively. In the next section, we use these assumptions about the source, in connection with the decoy state method and proof technique reviewed in Sec. II, to build up the security proof for this MDI QKD source.

C. Applying the proof technique

Given the model of the optical source, we can now connect it with the decoy state method reviewed in Appendix A and security proof technique from Sec. II A. To start, since the intentionally encoded states $\rho_{\text{enc}}^{k,i,x} \otimes \sigma_{\text{enc}}^{l,j,y}$ can still be treated as phase-randomized WCPs, and since the side-channel states $|\chi_x^i\rangle_{\text{leak}} \otimes |\zeta_y^j\rangle_{\text{leak}}$ carry no information about the decoy state intensity or random phase, we are able to use the decoy state method with only modifications to how we interpret the detection probabilities obtained by solving the linear programs.

The photon number distribution of the states $\rho_{\text{enc}}^{k,i,x} \otimes \sigma_{\text{enc}}^{l,j,y}$ follows the form from Eq. (A1); however, when considering the linear equations provided by the detection probabilities in Eq. (A2), $p_{\text{pass},m,n}^{i,j,x,y}$ now refers to the probability a round passes given that Alice sent the m -photon component of the state $\rho_{\text{enc}}^{k,i,x}$, that Bob sent the n -photon component of the state $\sigma_{\text{enc}}^{l,j,y}$, and that they together sent the leakage states $|\chi_x^i\rangle_{\text{leak}} \otimes |\zeta_y^j\rangle_{\text{leak}}$. Note that the state of the leakage light does not depend on the number of photons Alice and Bob sent, just on the polarization encoding choice, so we can still use m, n to label the variable, even though it does not strictly refer to Fock states anymore. Moreover, since the leakage states are independent of the intensity choice settings, $p_{\text{pass},m,n}^{i,j,x,y}$ remains independent of (k, l) . Solving the linear program in Appendix A, Alice and Bob retrieve, for each basis and bit setting (i, j, x, y) , bounds on the probabilities $p_{\text{pass},1,1}^{i,j,x,y}$ that Eve will announce a successful detection event given that they each sent a single photon in the encoded mode *along with* the associated side-channel state.

Interpreting $p_{\text{pass},1,1}^{i,j,x,y}$ as coming from both the single-photon components of the encoded mode *and* from the leakage light, means that the definition of the signal states in this protocol no longer refers just to the single-photon components of the encoded mode, as is the case for an ideal decoy state protocol. Connecting to Eq. (1), Alice and Bob's signal states are given by

$$|\psi_x^i \phi_x^j\rangle_{AB} = |\psi_x^i \phi_y^j\rangle_{\text{enc}} \otimes |\chi_x^i \zeta_y^j\rangle_{\text{leak}}, \quad (14)$$

where $|\psi_x^i \phi_x^j\rangle_{\text{enc}}$ refers to the single-photon components of the phase-randomized WCPs $\rho_{\text{enc}}^{k,i,x} \otimes \sigma_{\text{enc}}^{l,j,y}$. With these as the signal states, the detection probability constraints from Eq. (3)

employ the probabilities $p_{\text{pass},1,1}^{i,j,x,y}$ which come from the signal states, i.e., the leakage light and the single-photon component of the encoded mode. Additionally, the inner product constraints from Eq. (2) now include the inner products of the states of the leakage light; this reaffirms the versatility of the proof technique we are employing, since the constraints coming from the continuous-variable, time-dependent leakage light states can be coarse-grained down to their inner products, which form a finite-dimensional Gram matrix. The optimization variables (the elements of Eve's Gram G_E) and the objective function (the phase error rate) do not change; however, since the constraints will be affected by the presence of leakage light, the resulting key rate will certainly change.

With a model for the source, and how it connects to the decoy state method and the proof technique, we can now move to calculate the secret key rate for various scenarios and protocols.

IV. KEY RATE RESULTS

Having reviewed the main components required for the security proof technique in Sec. II, and having studied a case example of an MDI QKD source with a side channel in Sec. III, we now calculate key rates under different conditions. In Sec. IV A we provide the details of how our simulations were performed. Then in Sec. IV B we see how the key rate can change depending on the model chosen for the side channel, including the time-dependent state we derived in the previous section, highlighting the need for careful side-channel characterization. In Sec. IV C we explore a strategy for extracting higher key rates by sending states that would be redundant under ideal conditions but which help in the presence of a side channel, concluding that the three-state protocol and BB84, which yield equivalent key rates in the ideal case of no leakage light, have different key rates when a side channel is present. In Appendix E we examine two more strategies for boosting the key rate by using all mismatch statistics and by modifying the polarization angle of the states sent. These simulations should be used to better inform the choice of protocol when working with realistic sources like the one we are studying. While having considered only a specific source with a nontrivial side channel, we expect the conclusions drawn to be broadly applicable to any leaky source; namely, we emphasize the importance of side-channel characterization and determine which protocol parameters (e.g. number of states sent) can lead to key rate improvement.

A. Simulation details

Recall we have two types of constraints to calculate the key rate: the inner product of the initial states and the detection probabilities. Before detailing how we simulate these, we first comment on how different aspects of the source model affect these constraints.

To start, we observe that the single-photon component of the encoded mode is the only quantity that affects both the inner product and detection probability constraints. However, whether that single-photon state is directly sent from a single-photon source or is a component of a phase-randomized WCP is irrelevant to the inner product constraint; the type of source is relevant only to the detection statistics constraint, since

single photons undergoing a lossy channel will provide a different result at a threshold detector than phase-randomized WCPs.

Next, we assume the usage of gated detectors that would be timed to receive the encoded optical pulses. Hence, in our simulations of the detection events, we assume that the state of the leakage light has no impact on the overall observed detection statistics, since they are in temporal modes that are not picked up by the detectors and the already weak leakage light would be even less bright after the lossy channel. As a consequence, in the calculations we present, the side-channel state affects only the inner-product constraint. We note that the proof technique could easily accommodate the case of detection statistics being affected by the leakage light, as this would just be simulating different values of $Q_{k,l}^{i,j,x,y}$ in Eq. (A2).

Since different aspects of the source model affect constraints in nontrivial ways, to better understand the key rate resulting from the source described in Sec. III, we provide comparisons to other optical source models. Specifically, we consider the following:

(1) Single-photon vs phase-randomized WCP sources: when calculating the key rate for a given single-photon component and side-channel state, i.e., for fixed inner product constraints in Eq. (2), how much is the key rate affected by those signal states being used directly vs in a decoy state method?

(2) Sensitivity to the side-channel model: we assume in the detection simulations that the side-channel state has no impact on the observed outcomes, so the detection probability constraint in Eq. (3) remains fixed even if we change the model for the leakage light. In Sec. III B, we provided a model for the source which resulted in a time-dependent coherent state. Were we to change this model, how much does the key rate change?

In the sections that follow, we will consider these high-level choices of the model, in addition to varying more practically rooted parameters like the intensity of the leakage light, plus the number and choice of encoded states sent.

For the choice of side-channel model, we compare three different approaches to treating the state of the leakage light:

Model 1: full encoding information leaked. In this model, we assume the leakage light state is of the form:

$$|\chi_x^i\rangle_{\text{leak}} = \sqrt{\epsilon} |\text{vac}\rangle + \sqrt{1-\epsilon} |i, x\rangle \quad (15)$$

with $\langle i, x | i', x' \rangle = \delta_{i,i'} \delta_{x,x'}$. This model has been used in previous studies of QKD source side channels [9–11]. This model makes a relatively pessimistic assumption, since any nonvacuum component of leakage light provides full information, while we know, for example, that the single-photon component would not be able to unambiguously encode all possible basis and bit choices (i, x) .

Model 2: time-independent coherent state. In this model, we assume the leakage light state is of the form:

$$|\chi_x^i\rangle_{\text{leak}} = |\beta \cos \theta_x^i\rangle_H \otimes |\beta \sin \theta_x^i e^{i\phi_x^i}\rangle_V. \quad (16)$$

The angles (θ_x^i, ϕ_x^i) are chosen to coincide with the polarization angles of the encoded mode. This model is more realistic in that we know the leakage light, as laser light, is in a coherent state; however, it does not account for the time-dependent

nature of the polarization encoding in the leakage light, which has the opportunity to act to our advantage since not every instant provides Eve with full encoding information. Time-independent coherent state leakage light was considered in the context of Trojan horse attacks in [12].

Model 3: time-dependent coherent state (multiple temporal modes). This model assumes the state from Eq. (13). It is our most accurate model of the source side channel we introduced in Sec. III. The inner product between two general, time-dependent coherent states is given by

$$\langle \beta | \alpha \rangle = e^{-\frac{1}{2} \int dt [|\beta(t)|^2 + |\alpha(t)|^2 - 2\beta^*(t)\alpha(t)]}, \quad (17)$$

which we use to calculate the inner product of the side-channel states in Eq. (13).

While we have three different models for the leakage light, we can still associate each of them to a fixed leakage light intensity, $|\alpha|^2$. For Model 1, we can set $\epsilon = e^{-|\alpha|^2}$. In Model 2, we can set $\beta = \alpha$, and in Model 3 we can set $\int dt |\alpha(t)|^2 = |\alpha|^2$. This means all the models have the same vacuum probability, i.e., chance of sending no information to Eve, while the non-vacuum components carry varying amounts of information about the basis and bit values.

For the simulation of the detection statistics, in all our simulations we assume detection of a single Bell state using the detector setup from [2], with detector efficiency of 50%, dark count rates of 10^{-6} per pulse, and loss in fiber of 0.2 dB/km, with symmetric channel lengths from Alice and Bob to Charlie. To isolate the effect coming from the side channel, we do not assume any misalignment in the source, but this could easily be added to the detection simulations. When simulating the decoy state method, we have Alice and Bob employ constant intensities of 0.05, 0.1, and 0.6; however, an additional layer of optimization for the decoy state intensities is possible, using our phase error calculation as a subroutine. The detection outcome probabilities $Q_{k,l}^{i,j,x,y}$ for the phase-randomized WCPs were simulated using the method from [40]. All of our calculations are in the asymptotic limit of infinite key length, and in the limit as the sifting rate goes to 1.

B. Benefits of side-channel characterization

Our main interest is determining how the key rate is affected by the presence of the side channel. Here we investigate how the key rate changes depending on the model for the side-channel light, the intensity of the light, and on whether the encoded modes are sent as a perfect single-photon source or as phase-randomized WCPs. In these simulations, we assume that Alice and Bob prepare BB84 states $\frac{|H\rangle \pm |V\rangle}{\sqrt{2}}$ and $\frac{|H\rangle \pm i|V\rangle}{\sqrt{2}}$, i.e., there are no encoding flaws. These simulations serve as a test to see how robust the key rate calculation is to the model of the leakage light; unsurprisingly, the key rate is highly dependent on the state of the leakage light.

In Fig. 3(a) we plot the key rate as a function of Alice-Charlie distance, assuming a decoy state protocol, for the three models of leakage light. Additionally, we vary the intensity of the leakage light across several orders of magnitude; using the lowest intensity signals from [41] as an order-of-magnitude reference for highly attenuated light, this places realistic leak-

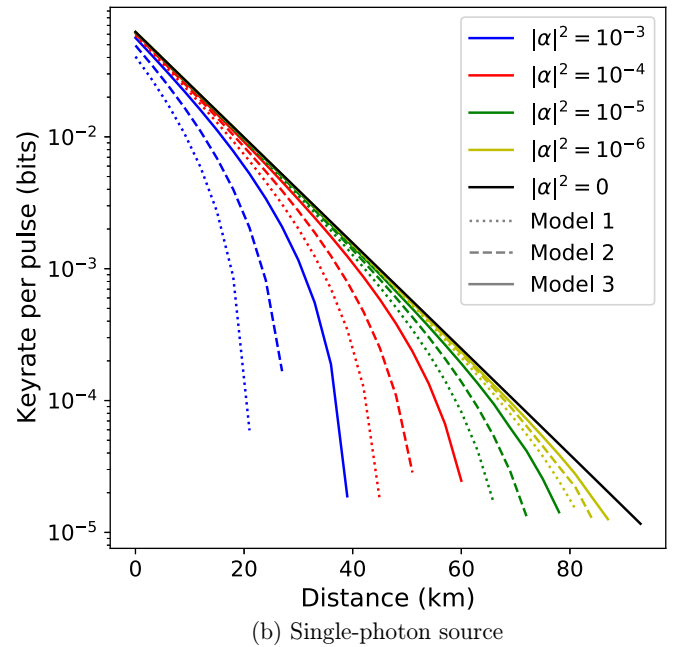
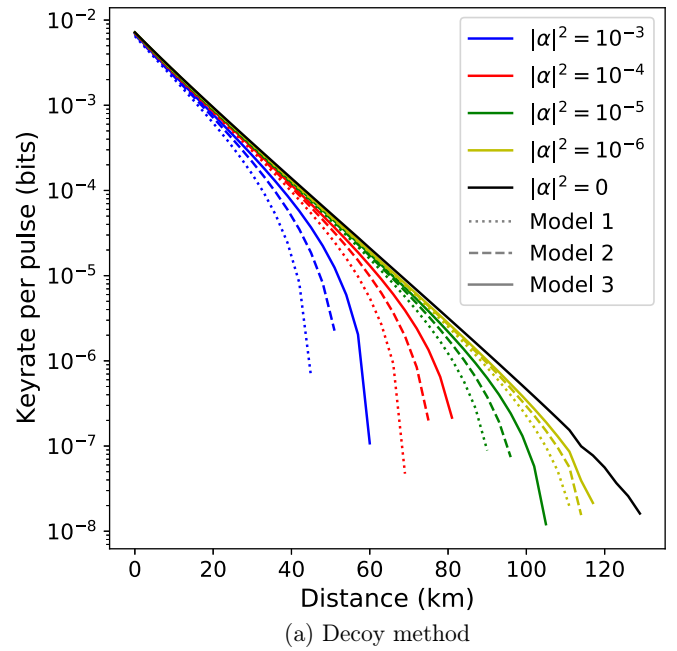


FIG. 3. Secret key rate as a function of Alice-Charlie distance for three different models of the leakage light. Model 1 corresponds to treating the leakage light as a superposition of vacuum (with amplitude $e^{-|\alpha|^2/2}$, same as a coherent state) and a state which leaks full encoding information. Model 2 corresponds to treating the leakage light as a coherent state $|\alpha_0 \cos \theta_x^i\rangle_H \otimes |\alpha_0 \sin \theta_x^i e^{i\phi_i}\rangle_V$ with the same polarization encoding parametrized by θ and ϕ as the signal state. Model 3 corresponds to treating the leakage light as a coherent state with total intensity $|\alpha|^2$, but with a time-dependent polarization, as given in Eq. (13). Across all models, $|\alpha|^2$ can be interpreted as the intensity of the leakage signal. For each model, we plot the key rate for various values of $|\alpha|^2$ which we indicate with different colours. In (a) we assume a decoy state protocol is used to characterize the single-photon detection events, while in (b) we assume that the encoded modes of the signal state are perfect single photons.

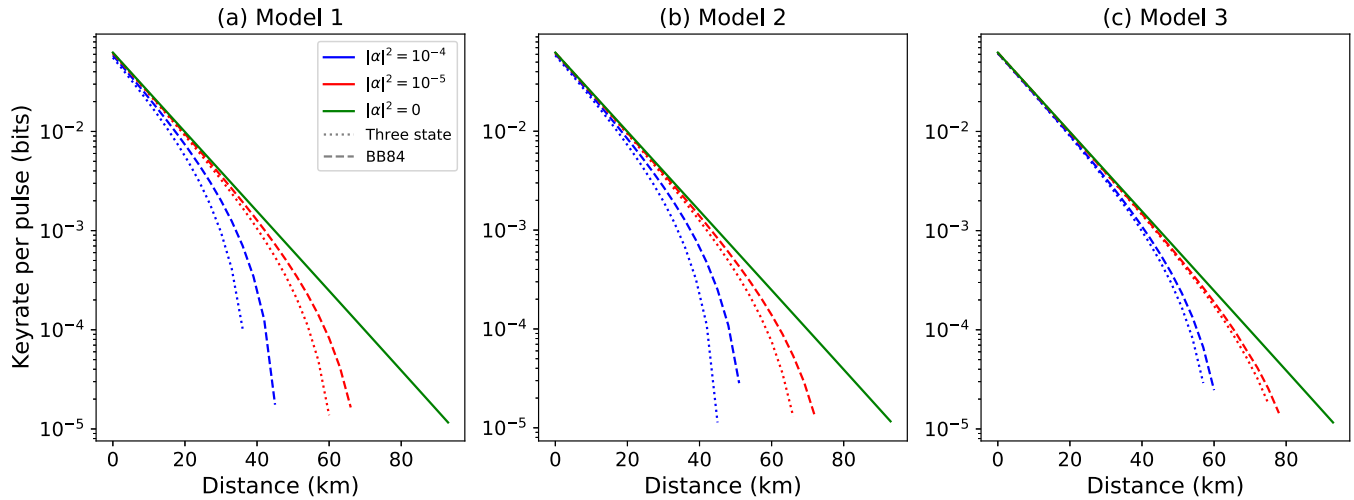


FIG. 4. Secret key rate as a function of Alice-Charlie distance assuming a single-photon source for the encoded mode states. Here we see the advantage of sending the four BB84 states instead of using the three-state protocol. This trend is true for different orders of magnitude of $|\alpha|^2$ and across all leakage state models, depicted in (a)–(c). This in contrast to the ideal case of $|\alpha|^2 = 0$, where BB84 and the three-state protocol yield the same key rates.

age light intensity somewhere on the order of 10^{-6} to 10^{-4} . From this plot, we see that the most significant boosts in key rate come from a hardware solution of minimizing the intensity of the light in the side channel; an order of magnitude reduction in intensity provides a greater improvement than refining the model of the leakage light state. However, there will likely always be some level of leakage light present between pulses. To mitigate this effect, it can be beneficial to carefully characterize the state of the side channel. We see an improvement in the key rate when moving from Models 1 through 3, in that order. This reflects the intuition that the nonvacuum components of the states in these models carry diminishing levels of information about the basis and bit choices. In MDI QKD, we require that Alice and Bob have complete characterization of their sources but no characterization of the detectors; thus, if they know the state of the side channel (or at the very least the pairwise inner products of all the initial states), it is straightforward to incorporate more information about the state by modifying the inner product constraints (a simple software solution), rather than making pessimistic assumptions about the leakage light, as in Model 1, resulting in lower key rates.

Since the model for and intensity of the side-channel light has no bearing on the observed detection statistics, the detection constraints used to produce all the key rate curves in the figure are the same. As extra confirmation that the observed improvements in the key rate due to changing the model of the side channel and the intensity of the side-channel light are independent of the observed detection statistics, in Fig. 3(b) we provide the same key rate calculations, but assume a single-photon source for the encoded mode. We notice qualitatively the exact same trends as when using the decoy method, as expected. We plot both the single-photon case and the decoy state case to demonstrate that the difference between the key rates for different models is due to the effect these models have when constraining the single-photon components, and that these differences persist even when we move to the realistic case of decoy states.

C. Sending seemingly redundant states helps

It is known that with only three out of the four BB84 states and using all the detection statistics, that one can produce the same key rate as using all the BB84 states [3]. Here we are interested in whether the same conclusion extends to the case of when the source has a side channel. While the fourth BB84 state is redundant in the case that there is no leakage light, when a side channel is present, the extra state can help characterize Eve's attack on the leakage light. We certainly would not expect the key rate to decrease by sending an extra state, as the extra state will only provide additional inner product and detection constraints to those already provided by the other three states.

In Fig. 4 we plot the key rate for a single-photon source, examining all three models of leakage light, and a couple different intensities. Across all models and intensities (except for $|\alpha|^2 = 0$) there is an increase in the key rate when all four BB84 states are used as opposed to only three. In Fig. 5 we plot the key rates again, this time assuming a decoy state protocol. In this case, the divergence between using three or four states is even more pronounced. Even the $|\alpha|^2 = 0$ case observes a boost in the key rate, since the detection statistic constraints in Eq. (4) are inequalities when using the decoy state method, so the extra detection statistics from the fourth state are useful in this case. The key rate boost achieved from switching from three to four states is so pronounced that it can even do better than decreasing the intensity of the leakage light: using four states with a leakage light intensity of 10^{-4} provides a higher key rate than using three states with a leakage intensity of 10^{-5} .

The takeaway message from this analysis is that the three-state protocol is not as practical as BB84 in the presence of source side channels. The extra resource savings of only having to use three states is undone by the loss of useful constraints that increase the key rate. We also simulated sending five and six states in the same plane of the Bloch sphere as the BB84 states to see if this provided even better key rates, but the key rate appeared to saturate with sending four states.

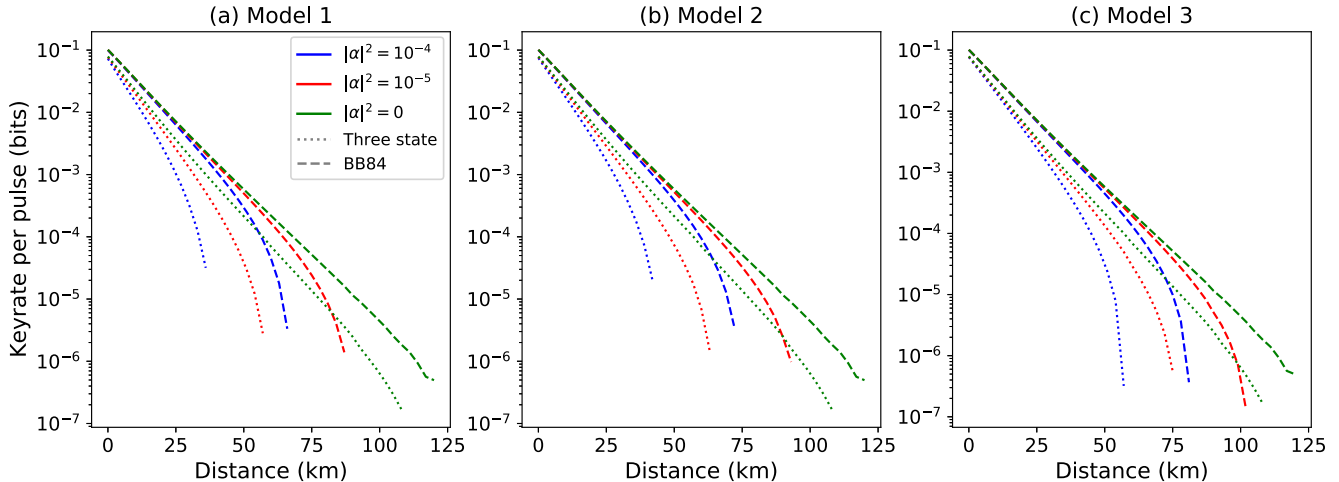


FIG. 5. Secret key rate as a function of Alice-Charlie distance assuming a phase-randomized WCP and decoy state method. Like in Fig. 4, we see the advantage of sending the four BB84 states instead of using the three-state protocol. Comparing the dashed green line to the dotted green line, we see that even in the ideal case of $|\alpha|^2 = 0$, BB84 outperforms the three-state protocol, since the detection probabilities only offer inequality constraints, meaning the extra fourth state does offer extra constraint to increase the key rate. When the side channel is present, we also see that adding an extra state in BB84 can go so far as to achieve a higher key rate than using the three-state protocol with an order of magnitude lower leakage light intensity.

Certainly sending additional states outside of this plane would increase the key rate, as expected from the six-state [42] or tilted four-state protocols [3], but this would require additional polarization modulation in the source, whereas it is easier to only vary the angle along one great circle of the Bloch sphere.

V. CONCLUSION

In this work, we have examined the problem of source side channels in MDI QKD. We reviewed the decoy state method and a recent, versatile proof technique based on semidefinite programming which allows for information about the state of the side channel to be incorporated into the key rate calculation. With this in hand, we examined a case example of a common MDI QKD source that employs a Faraday mirror for polarization stabilization. For this source, we identified a nontrivial, time-dependent side channel due to leakage light between encoded optical pulses, provided a quantum optical model of the output, and linked the components of the source model to the security proof techniques. We then examined multiple protocol scenarios to understand strategies for improving the secret key rate under practical circumstances.

We identified how the key rate calculation is affected by the information provided as constraints to the security proof. Most importantly, we saw how the model for the state of the leakage light can significantly impact the key rate, reaffirming that in MDI QKD security is derived in part from knowledge of the initial states sent by Alice and Bob, including any side-channel states. It is clear from our results that in practice one must carefully characterize side channels, the reward of this work being higher key rates that come from not needing to take overly pessimistic assumptions of how much information is being leaked to the eavesdropper. On top of the importance of using the best available model for the side channel, we found that in the presence of state-preparation flaws, Alice

and Bob benefit from using all information at their disposal for the key rate calculation, i.e., all detection statistics, and all initial state information, rather than discarding cases when their basis choices do not match.

Having models for the state of the leakage light allowed us to develop concrete hardware strategies for mitigating the presence of the side channel. Besides the obvious hardware improvement of simply suppressing the leakage light, two other physically implementable strategies emerged for when leakage light is present: first, although the three-state protocol promises the same key rates as BB84, when leakage light is present, Alice and Bob can get better key rates by sending all four BB84 states, as the statistics from the normally redundant fourth state actually help to better constrain Eve's attack on the side channel. Second, while the choice of which test state to send from the Bloch sphere typically does not matter, here we find that in the presence of leakage light, some test states provide better key rates than others, indicating the advantage of optimizing which states to send as a function of distance. Even though we considered a representative case example, we expect that the strategies we developed to mitigate the side channel to be widely applicable to other leaky sources.

While this work examined strategies for treating source side channels in MDI QKD, the source we considered had the advantage of not leaking information about the intensity setting choice and random phase value of the decoy state protocol, meaning we were able to use the decoy state method with only modifications to how we interpret the output of the linear programs in the security proof. An open problem is how to simultaneously mitigate more general source side channels that leak information about both the encoding information, as we investigated, and the decoy state method intensity and phase parameters. It would be worthwhile to investigate merging the analysis presented here with the proof technique from [8] for treating intensity and phase information leakage.

ACKNOWLEDGMENTS

We are grateful for discussions with Thomas Van Himbeek, Ignatius William Primaatmaja, and Emilien Lavie, and especially thankful to Wenyuan Wang for extensive discussions about decoy states. J.E.B is supported by an Ontario Graduate Scholarship. We acknowledge financial support from NSERC, MITACS, CFI, ORF, the Royal Bank of Canada, Huawei Technology Canada, CRCEF, and the University of Hong Kong start-up funding.

APPENDIX A: REVIEW OF THE DECOY STATE METHOD FOR MDI QKD

When Alice and Bob encode their secret key in a single-photon degree of freedom, photon number splitting attacks are a method for an eavesdropper to exploit multiphoton output of the optical source to learn the secret key [37]. As a consequence, only single-photon detection events are usable to characterize the amount of information the eavesdropper has about the key; however, multiphoton events can still be used to characterize the correctness of the key. The decoy state method allows Alice and Bob to characterize the photon number statistics of the eavesdropper-controlled channel, and in turn bound security based on the detection events that arose from the single-photon components of the source's optical output [43].

Practically, the decoy state method for MDI QKD with polarization encoding consists of Alice and Bob each preparing phase-randomized WCPs with varying intensities [40]. Each pulse is polarized according to the protocol, e.g., BB84 [2] or three-state [3]. In this case, we can write the photon number distribution of Alice and Bob's states as

$$p(m, n|k, l) = \frac{e^{-(\mu_k + \nu_l)} \mu_k^m \nu_l^n}{m!n!}, \quad (\text{A1})$$

where k (l) refers to Alice's (Bob's) optical intensity setting μ_k (ν_l). This assumes that the intensity settings are completely independent of the polarization basis and bit setting choices. If they each use N intensity settings (typically three is sufficient), then for given basis (i, j) and bit (x, y) choices, they have N^2 linear equations for the detection probabilities as a function of photon number:

$$Q_{k,l}^{i,j,x,y} = \sum_{m,n} \frac{e^{-(\mu_k + \nu_l)} \mu_k^m \nu_l^n}{m!n!} p_{\text{pass},m,n}^{i,j,x,y}. \quad (\text{A2})$$

Here $Q_{k,l}^{i,j,x,y}$ is the observed probability of Eve announcing that a round passed given that Alice and Bob chose intensity settings (k, l) along with basis and bit choices (i, j, x, y) . $p_{\text{pass},m,n}^{i,j,x,y}$ is the probability that the round passes due Alice (Bob) sending m (n) photons, and given basis and bit choices (i, j, x, y) . Since we assume that the intensity setting choices are independent of the basis and bit choices, and that phase randomization of the WCP is perfect, $p_{\text{pass},m,n}^{i,j,x,y}$ is independent of (k, l) . These N^2 linear equations can be used in a linear program to determine upper and lower bounds on all the detection probabilities due to single-photon components of the optical output $p_{\text{pass},1,1}^{i,j,x,y}$. This allows us to bound the relevant detection statistics for computing security.

To construct a linear program to calculate upper and lower bounds on the single-photon detection probabilities of a decoy state MDI QKD protocol [40], we first identify the variables of the optimization as $p_{\text{pass},m,n}^{i,j,x,y}$. To establish a lower (upper) bound on $p_{\text{pass},1,1}^{i,j,x,y}$ given these constraints, we solve the linear program to find the minimum (maximum) possible value of $p_{\text{pass},1,1}^{i,j,x,y}$ consistent with the constraints. If Alice and Bob each have n_A and n_B basis choice settings, each basis choice associated with two bit choices, we repeat the process of finding lower and upper bounds for all $n_A \times n_B \times 2 \times 2$ combinations of (i, j, x, y) .

Since there are in principle infinitely many $p_{\text{pass},m,n}^{i,j,x,y}$, for a practical linear program, we impose a cutoff photon number N_{max} . In that case, the N^2 linear equality constraints become $2N^2$ linear inequality constraints. The first N^2 constraints are

$$Q_{k,l}^{i,j,x,y} \geq \sum_{m,n=0}^{N_{\text{max}}} \frac{e^{-(\mu_k + \nu_l)} \mu_k^m \nu_l^n}{m!n!} p_{\text{pass},m,n}^{i,j,x,y}, \quad (\text{A3})$$

stemming from the fact that summing up to the cutoff will yield a value less than the total detection probability. For the next N^2 constraints, we find

$$1 - \sum_{m,n=0}^{N_{\text{max}}} \frac{e^{-(\mu_k + \nu_l)} \mu_k^m \nu_l^n}{m!n!} \geq \sum_{m,n=N_{\text{max}}+1}^{\infty} \frac{e^{-(\mu_k + \nu_l)} \mu_k^m \nu_l^n}{m!n!} p_{\text{pass},m,n}^{i,j,x,y}, \quad (\text{A4})$$

which means we can provide the constraints:

$$\sum_{m,n=0}^{N_{\text{max}}} \frac{e^{-(\mu_k + \nu_l)} \mu_k^m \nu_l^n}{m!n!} p_{\text{pass},m,n}^{i,j,x,y} \geq Q_{k,l}^{i,j,x,y} + \sum_{m,n=0}^{N_{\text{max}}} \frac{e^{-(\mu_k + \nu_l)} \mu_k^m \nu_l^n}{m!n!} - 1. \quad (\text{A5})$$

In practice for our calculations in the main text, we found an N_{max} of 10 photons was sufficient to provide good upper and lower bounds on $p_{\text{pass},1,1}^{i,j,x,y}$ while not taking too long to compute.

Using the decoy method, a lower bound on the secret key rate in an MDI QKD protocol is provided by [40]

$$R \geq p_{\text{pass},1,1}^{0,0} [1 - h_2(e_{ph,1,1})] - Q_{N,N}^{0,0} h_2(E_{bit}), \quad (\text{A6})$$

where $h_2(\cdot)$ is the binary entropy function. $Q_{N,N}^{0,0}$ is the detection probability of outcomes that generate raw key, and is given by

$$Q_{N,N}^{0,0} = \sum_{x,y} Q_{N,N}^{0,0,x,y}, \quad (\text{A7})$$

where we choose, without loss of generality, $(i, j) = (0, 0)$ to be the key generation basis and $(k, l) = (N, N)$ to be the key generation intensities. E_{bit} is the bit error rate of the raw key, given by

$$E_{bit} = \frac{\sum_{x \neq y} Q_{N,N}^{0,0,x,y}}{Q_{N,N}^{0,0}}. \quad (\text{A8})$$

$p_{\text{pass},1,1}^{0,0}$ is the detection probability due to the single-photon components of Alice and Bob's optical output:

$$p_{\text{pass},1,1}^{0,0} = \sum_{x,y} p_{\text{pass},1,1}^{0,0,x,y}. \quad (\text{A9})$$

Finally, $e_{ph,1,1}$ is the phase error rate of the protocol which we will more precisely define in the next Appendix. Briefly, were we to consider a virtual picture of the protocol in which the single-photon components of the optical output are entangled with qubits kept in Alice and Bob's labs, the phase error is the probability those qubits end up in a target Bell state up to a phase error. Like the single-photon detection probabilities, it is not a directly observable quantity of the protocol and must be bounded. Were Alice and Bob to be able to perfectly prepare eigenstates of the conjugate basis to the key generation basis, then $e_{ph,1,1}$ can also be bounded with a simple linear program [40]. When the sources have preparation flaws amounting to constant polarization offsets, a series of linear programs are required (see Appendix A of [4]). In the case that the sources have preparation flaws and have a side channel, we can employ the more recent technique [12] for bounding the phase error rate that employs semidefinite programming; that technique is reviewed in Sec. II A.

APPENDIX B: DERIVATION OF EQ. (7)

Here we go through the derivation of the expression for the phase error rate in terms of the elements of the Gramian matrix. We wish to show

$$\begin{aligned} e_{ph} &= \frac{\langle \Gamma | (\mathbb{M}_{\bar{A}\bar{B}}^- \otimes |P\rangle \langle P|_Z) | \Gamma \rangle_{\bar{A}\bar{B}EZ}}{\langle \Gamma | (|P\rangle \langle P|_Z) | \Gamma \rangle_{\bar{A}\bar{B}EZ}} \\ &= \frac{1}{2} - \frac{\text{Re}(\langle e_{0,0,P}^{0,0} | e_{1,1,P}^{0,0} \rangle_E + \langle e_{0,1,P}^{0,0} | e_{1,0,P}^{0,0} \rangle_E)}{\sum_{x,y} p_{\text{pass}}^{0,0,x,y}}. \end{aligned} \quad (\text{B1})$$

We start with the denominator:

$$\begin{aligned} \langle (|P\rangle \langle P|_Z) | \Gamma \rangle_{\bar{A}\bar{B}EZ} &= \sum_{x,y=0}^1 |x,y\rangle_{\bar{A}\bar{B}} \langle e_{x,y,P}^{0,0} \rangle_E \\ &\Rightarrow \langle \Gamma | (|P\rangle \langle P|_Z) | \Gamma \rangle_{\bar{A}\bar{B}EZ} \\ &= \sum_{x,y,x',y'=0}^1 \langle x,y | x',y' \rangle_{\bar{A}\bar{B}} \langle e_{x,y,P}^{0,0} | e_{x',y',P}^{0,0} \rangle_E \\ &= \sum_{x,y=0}^1 \langle e_{x,y,P}^{0,0} | e_{x,y,P}^{0,0} \rangle_E \\ &= \sum_{x,y} p_{\text{pass}}^{0,0,x,y}, \end{aligned} \quad (\text{B2})$$

where the last equality follows from Eq. (3).

Next we look at

$$\begin{aligned} \mathbb{M}_{\bar{A}\bar{B}}^- &= (|\Phi^-\rangle \langle \Phi^-| + |\Psi^-\rangle \langle \Psi^-|)_{\bar{A}\bar{B}} \\ &= \frac{1}{2} [\mathbb{1} - (|0,0\rangle \langle 1,1| + \text{H.c.} + |0,1\rangle \langle 1,0| + \text{H.c.})]_{\bar{A}\bar{B}}. \end{aligned} \quad (\text{B3})$$

The identity term will simply yield the same expression as the denominator we have already calculated, leading to the $\frac{1}{2}$ term

in Eq. (7). The other terms can be calculated as follows:

$$\begin{aligned} &\langle \Gamma | (|0,0\rangle \langle 1,1|_{\bar{A}\bar{B}} + \text{H.c.}) \otimes (|P\rangle \langle P|_Z) | \Gamma \rangle_{\bar{A}\bar{B}EZ} \\ &= \sum_{x,y,x',y'=0}^1 \langle x,y | 0,0 \rangle \langle 1,1 | x',y' \rangle_{\bar{A}\bar{B}} \langle e_{x,y,P}^{0,0} | e_{x',y',P}^{0,0} \rangle_E + \text{c.c.} \\ &= \langle e_{0,0,P}^{0,0} | e_{1,1,P}^{0,0} \rangle_E + \text{c.c.} \\ &= 2\text{Re}(\langle e_{0,0,P}^{0,0} | e_{1,1,P}^{0,0} \rangle_E). \end{aligned} \quad (\text{B4})$$

Similarly,

$$\begin{aligned} &\langle \Gamma | (|0,1\rangle \langle 1,0|_{\bar{A}\bar{B}} + \text{H.c.}) \otimes (|P\rangle \langle P|_Z) | \Gamma \rangle_{\bar{A}\bar{B}EZ} \\ &= 2\text{Re}(\langle e_{0,1,P}^{0,0} | e_{1,0,P}^{0,0} \rangle_E). \end{aligned} \quad (\text{B5})$$

These two terms yield the numerator of the second term in Eq. (7).

APPENDIX C: COMPARISON TO THE PROOF TECHNIQUE FROM [11]

Here we will compare the proof technique we are using to the technique from [11]. We will show that [11] relaxes the SDP inherent to optimizing the phase error rate to a linear program, which we would expect to give equal or lower bounds on the key rate than computing the full SDP. For simplicity, we will consider a protocol where when Alice and Bob choose the Z basis, they perfectly prepare qubit states $|0\rangle, |1\rangle$, but we allow for their test states to have leakage components outside of the qubit subspace spanned by $\{|0\rangle, |1\rangle\}$. The following comparison can be generalized in a straightforward manner to arbitrary initial states.

Let U be the unitary that takes $|\psi_x^i \phi_y^j\rangle_{A,B} \rightarrow \sum_z |e_{x,y,z}^{i,j}\rangle$. Thus, for this case of initial states, the phase error rate would be given by

$$e_{ph} = \frac{\text{Tr}[|P\rangle \langle P|_Z U \left(\frac{\mathbb{1} + \sigma_X \otimes \sigma_X}{2} \right)_{A,B} U^\dagger]}{\sum_{x,y} p_{\text{pass}}^{0,0,x,y}}, \quad (\text{C1})$$

where $\sigma_m, m = I, X, Y, Z$ refer to Pauli operators in the qubit space spanned by $\{|0\rangle, |1\rangle\}$. Following [3, 11], e_{ph} can be decomposed in terms of the transmission rates of the Pauli operators $q_{\text{pass}|i,j} = \text{Tr}(|P\rangle \langle P|_Z U \sigma_i \otimes \sigma_j U^\dagger)$, since the Pauli operators form a basis for any operator. Were $\epsilon = 1$, then we could use the states Alice and Bob send to exactly solve for $q_{\text{pass}|i,j}$ (assuming their test states are some superposition of $\{|0\rangle, |1\rangle\}$). However, when their signal states have a leakage component, we cannot exactly constrain these quantities; in [11] $q_{\text{pass}|i,j}$ form the variables of a linear program that are optimized to determine a lower bound on e_{ph} .

First, we write Alice and Bob's signal states as linear combinations of states in a two-qubit space, and a space orthogonal to it (the leakage space), just as in Eq. (1) of [11]:

$$|\psi_x^i \phi_y^j\rangle_{A,B} = a_{x,y}^{i,j} |\tilde{\psi}_x^i \tilde{\phi}_y^j\rangle_{A,B} + b_{x,y}^{i,j} |\tilde{\psi}_x^i \tilde{\phi}_y^{j\perp}\rangle_{A,B}, \quad (\text{C2})$$

where the orthogonality of the two-qubit and leakage space means $\langle \cdot | \cdot^\perp \rangle_{A,B} = 0$.

Next, the detection probabilities provide the constraint

$$p_{\text{pass}}^{i,j,x,y} = \text{Tr}(|P\rangle \langle P|_Z U |\psi_x^i \phi_y^j\rangle \langle \psi_x^i \phi_y^j|_{A,B} U^\dagger). \quad (\text{C3})$$

Using the decomposition from Eq. (C2), we find this is also equal to

$$\begin{aligned}
p_{\text{pass}}^{i,j,x,y} &= |a_{x,y}^{i,j}|^2 \text{Tr}[|P\rangle\langle P|_Z U |\tilde{\psi}_x^i \tilde{\phi}_y^j\rangle\langle \tilde{\psi}_x^i \tilde{\phi}_y^j|_{A,B} U^\dagger] \\
&+ \text{Tr}[|P\rangle\langle P|_Z U (a_{x,y}^{i,j} b_{x,y}^{i,j*} |\tilde{\psi}_x^i \tilde{\phi}_y^j\rangle\langle \tilde{\psi}_x^i \tilde{\phi}_y^{j\perp}| \\
&+ a_{x,y}^{i,j*} b_{x,y}^{i,j} |\tilde{\psi}_x^i \tilde{\phi}_y^{j\perp}\rangle\langle \tilde{\psi}_x^i \tilde{\phi}_y^j| \\
&+ |b_{x,y}^{i,j}|^2 |\tilde{\psi}_x^i \tilde{\phi}_y^{j\perp}\rangle\langle \tilde{\psi}_x^i \tilde{\phi}_y^{j\perp}|)_{A,B} U^\dagger], \quad (\text{C4})
\end{aligned}$$

which coincides with the MDI QKD version of Eq. (19) from [11]. Since the operator $|\tilde{\psi}_x^i \tilde{\phi}_y^j\rangle\langle \tilde{\psi}_x^i \tilde{\phi}_y^j|$ lives in the two-qubit subspace, it too can be written in terms of the Pauli operators $\sigma_m \otimes \sigma_n$, meaning the first trace term in Eq. (C4) can be written in terms of $q_{\text{pass}|i,j}$. Because of the second trace term, we cannot solve for them exactly as is done in the loss-tolerant proof technique [3].

Rather than solving the semidefinite program for e_{ph} using the linear equality constraints provided by the detection probabilities, [11] considers a relaxation to a linear program. Specifically, the second trace term in Eq. (C4) can be bounded above and below by the maximum and minimum eigenvalues of the matrix:

$$M_{x,y}^{i,j} = \begin{pmatrix} 0 & a_{x,y}^{i,j} b_{x,y}^{i,j*} \\ a_{x,y}^{i,j*} b_{x,y}^{i,j} & |b_{x,y}^{i,j}|^2 \end{pmatrix}. \quad (\text{C5})$$

Thus, (C4) leads to inequalities linear in $q_{\text{pass}|i,j}$, which can be used as constraints in a linear program to find an upper bound for e_{ph} . However, because the exact equality constraints coming from the detection probabilities have been relaxed using the maximum and minimum eigenvalues of $M_{x,y}^{i,j}$, we would expect that this would lead to a greater upper bound on e_{ph} (and hence a weaker lower bound on the key rate) than if the exact constraints were kept, as they would be in the numerical approach from [12] that we reviewed in Sec. II A.

Note that the proof approach we have used in this paper does away with needing to frame the calculation of e_{ph} in terms of $q_{\text{pass}|i,j}$ (even though we could, in principle, do so since they are linear functions of the elements of Eve's Gram matrix); after all, since the signal states are no longer qubits, we need not make the distinction between a qubit subspace and the leakage space, since Eve's operation can blend these two spaces. Instead, given that the phase error can be expressed in terms of the elements of a positive semidefinite matrix associated with Eve's information, and given that we have linear equality constraints on this matrix, the phase error can be maximized directly with a simple SDP, rather than relaxing to a linear program.

As an example to demonstrate the superiority of the SDP method over the method from [11], we consider a toy example of the three-state protocol with a single-photon source, for which Alice and Bob prepare a leaky third state,

$$|+\rangle_{\text{enc}} (\sqrt{\epsilon} |vac\rangle_{\text{leak}} + \sqrt{1-\epsilon} |1\rangle_{\text{leak}}), \quad (\text{C6})$$

as opposed to the ideal $|+\rangle_{\text{enc}}$. We assume a detection efficiency of 1 and a dark count rate of 10^{-6} . In Fig. 6 we plot the key rates calculated using the SDP method we reviewed in Sec. II A and using the method from [11]. We find that across values of ϵ , the SDP method performs much better.

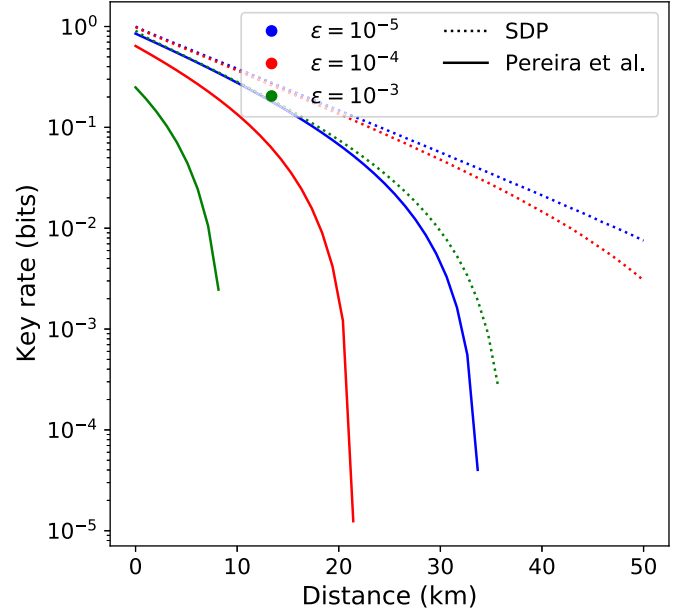


FIG. 6. Key rate vs distance for various values of ϵ , calculated using the SDP method reviewed in Sec. II A and the method from [11]. We see the advantage of the SDP approach over the relaxation to a linear program, as done in [11].

APPENDIX D: DERIVATION OF FIG. 2

Here we derive the fractional phase change applied to the leakage light as a function of time, as shown in Fig. 2.

Refer to the experimental setup shown in Fig. 1, specifically the polarization modulation unit. First, optical pulses travel forward through the PM for the purpose of polarization modulation. Simultaneously, voltage pulses overlapping in time with the optical pulses are sent into the PM, propagating in the same direction as the optical pulses. The PM is designed such that the optical and voltage pulses travel through the PM at the same speed [44]. The voltage is what enables a phase change and therefore a polarization change. Since the leakage light is not meant to encode information, voltage is not sent through the PM as this light travels through the first time.

However, when traveling back through the PM after reflection from the Faraday mirror, the leakage light will inevitably collide temporally with a voltage pulse that is traveling in the opposite direction along with an optical pulse it is intended to modulate.

The overall phase modulation experienced by a temporal slice of light after traveling through the PM can be expressed as

$$\phi = K \int_0^L V(z) dz. \quad (\text{D1})$$

Here L represents the length of the PM and $V(z)$ represents the applied voltage overlapping with the slice of light. K is simply a proportionality constant. When light is traveling in the same direction as the voltage wave through the phase modulator, Eq. (D1) reduces to $K \times V \times L$. This occurs due to the voltage, which is moving at the same speed as the light, being a constant along the length of the PM.

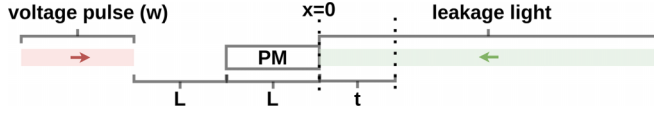


FIG. 7. An illustration depicting a moment in time as a voltage pulse approaches the phase modulator (PM).

In our case, we are also dealing with leakage light that is traveling in the opposite direction. We will use L to refer to the PM length and w to refer to the width of the square voltage pulses sent to the PM. Given these parameters, we can determine the phase change experienced by the leakage light as follows:

(1) Refer to Fig. 7. We will use this moment in time as our starting point. First we will create a coordinate system by defining $x = 0$ to be the right hand edge of the phase modulator. We can parametrize a slice of leakage light with t , the time it crosses the point $x = 0$.

(2) At the point in time shown in Fig. 7, we can define the voltage pulse as $A(x) = H(x + 2L + w) - H(x + 2L)$ (a square pulse) and the leakage light as $B(x) = H(x)$ where H refers to the Heaviside step function.

(3) Now, notice that the movement in time of the voltage pulse and leakage light can also be incorporated into these functions. After τ ps, the function defining the voltage pulse will become $A(x - \tau)$ while the function defining the leakage light will become $B(x + \tau)$.

(4) Notice that $A(x - \tau) \times B(x + \tau)$ represents the overlap between the voltage and leakage light at position x and time τ . It has a value of 1 if there is an overlap and a value of 0 if there is no overlap.

(5) Now, suppose we want to calculate the amount of time for which the slice t experiences an overlap within the phase modulator. We need to integrate $A(x - \tau) \times B(x + \tau)$ from $\tau = t$ to $\tau = t + L$. In other words, we need to integrate the overlap function over the values of τ for which the slice at position t is inside the phase modulator.

(6) The slice at position t has an x position of $t - \tau$ at time τ . Substitute this into the integral for x .

(7) The resulting integral is as follows:

$$\int_t^{t+L} A(t - 2\tau) \times B(t) d\tau. \quad (\text{D2})$$

This integral represents the amount of time the slice t is in contact with a voltage pulse within the phase modulator. Recall that the optical pulse which is traveling along with (in the same direction as) this voltage pulse would be in contact along the entire length of the phase modulator (L). Therefore, the phase change experienced by slice t is

$$\frac{1}{L} \int_t^{t+L} A(t - 2\tau) \times B(t) d\tau \quad (\text{D3})$$

when written as a fraction of the phase change experienced by the pulse. The solution to this integral is shown in Fig. 8. In our particular experimental setup, $L = 150$ ps and $w = 200$ ps. The value of the integral for these parameter values is plotted in Fig. 2. The maximal fractional phase change is $\frac{2}{3}$.

APPENDIX E: ADDITIONAL KEY RATE RESULTS

Here we present additional exploration of strategies that can be used to increase the key rate in the presence of a source side channel.

1. Basis mismatch constraints

We are interested in knowing whether any advantage can be gained by using all the detection statistics and all the initial state inner products, including when Alice and Bob's bases do not match, as opposed to simply using the cases when the basis choices match ($i = j$). We observed that when Alice and Bob prepare the BB84 states perfectly, using the basis mismatch statistics and inner products did not produce an increase in the key rate, even in the presence of a side channel. With perfect state preparation, we know that the conjugate basis statistics alone are strict enough constraints to provide the phase error when there is no leakage light, and we confirm numerically that this extends to the case when leakage light is present.

However, we know that when Alice and Bob have a preparation flaw for their states, i.e., a constant offset angle on the Bloch sphere, the mismatch statistics can help better characterize the key rate [3]. Since the Bloch sphere angle affects the associated side-channel state, the inner products, and the detection statistics, it is more difficult to predict how the key rate will respond to a preparation flaw, and whether using full or partial detection statistics in the SDP constraints benefits

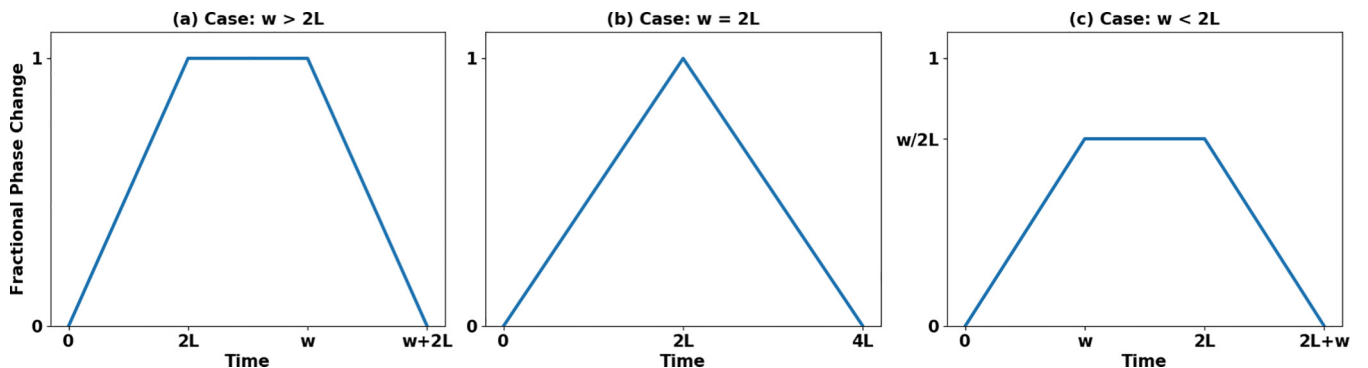


FIG. 8. The various solutions for the integral in Eq. (D3).

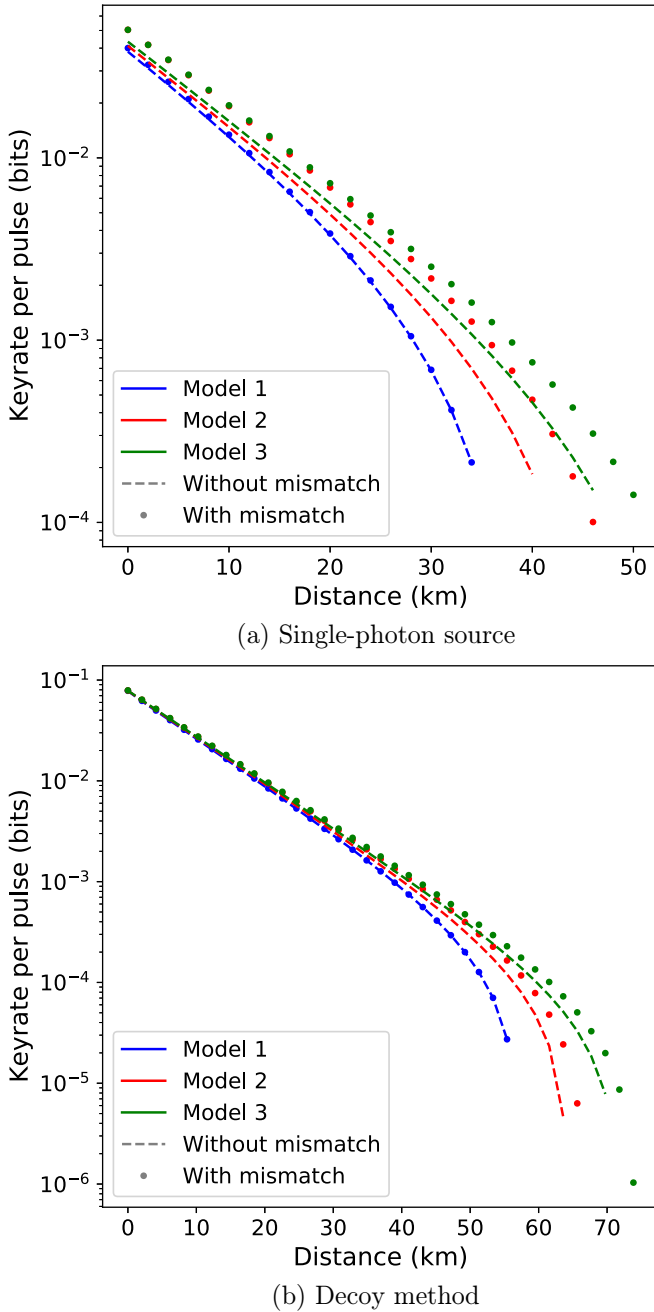


FIG. 9. Secret key rate as a function of Alice-Charlie distance for Models 1–3 of the leakage light. Here we investigate whether using all initial state inner products and detection statistics, as opposed to just the cases when Alice and Bob choose the same basis, benefit the key rate. We consider the case of BB84 with a preparation flaw, and a side channel with $|\alpha|^2 = 10^{-4}$, for both (a) a single-photon source, and (b) the decoy state method. For (b), Alice and Bob each uses four decoy intensities. We observe that for Models 2 and 3, the key rate benefits from considering all inner products and detection statistics available.

the key rate. For these simulations, we use the preparation flaw model from Appendix D of [4], with the Bloch sphere offset angle parameter $\delta = 0.1$. In this case we fix the leakage light intensity to $|\alpha|^2 = 10^{-4}$.

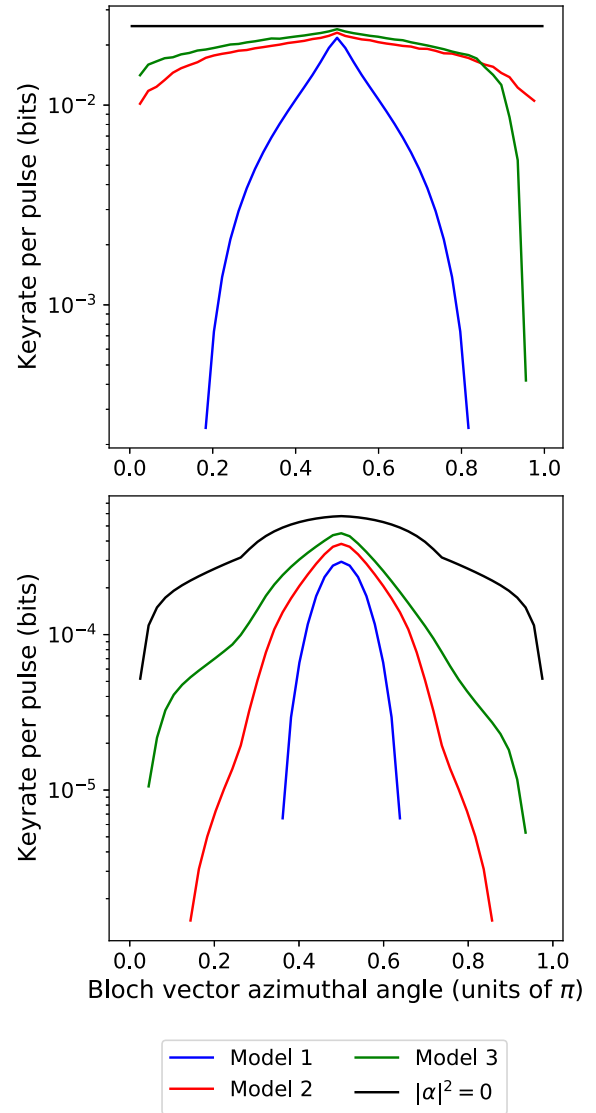


FIG. 10. Key rate vs azimuthal angle of the test states, for the case of no leakage light, and for leakage light with intensity $|\alpha|^2 = 10^{-4}$ treated with Models 1–3. The top (bottom) figure provides results for a single-photon source (a decoy state method) at a distance of 10 km (50 km). While the choice of test state is less relevant for the case of no leakage light, it can significantly decrease the key rate in the presence of leakage light, prompting the need to optimize which test states are used at a given distance.

In Fig. 9(a) we plot the key rate assuming a single-photon source for the encoded mode. For Model 1 we barely see any increase in the key rate when using full vs partial detection statistics; this makes sense, since the nonvacuum component of the side-channel state leaks full encoding information, independent of Bloch sphere angle. For Models 2 and 3 we observe a boost in the key rate when using full detection statistics and inner products as constraints. This indicates that when one has a preparation flaw, and the side-channel state depends on the preparation flaw, it is best to use all information available from the detection statistics and initial state inner products.

In Fig. 9(b) we consider the same situation but with a decoy state protocol. For this scenario, we add a fourth decoy

with vacuum intensity, and observe an increase in the key rate when using full detection statistics and initial inner products as constraints in Models 2 and 3. Like before, we do not observe an increase in the key rate for Model 1. When we only considered three decoy intensities, we did not observe a meaningful increase in the key rate, likely because the three decoy intensities did not allow for tight enough constraints on the single-photon detection statistics, so adding more detection statistics as constraints did not help since the constraints were too loose.

2. Choice of test states matters

Another example of divergence between ideal sources and sources with side channels occurs in the choice of which test states to send. In the ideal case, if Alice and Bob prepare two orthogonal polarization states, they need only send one other state to achieve the same key rate as BB84 [3]; the location of that state on the Bloch sphere does not matter (as long as it is not the same state as the first two). Here we are interested to see whether this changes in the presence of a source side channel.

To study this problem, we fix the channel distance, a leakage light intensity of $|\alpha|^2 = 10^{-4}$, fix Alice and Bob to send encoded single-photon components $\frac{|H\rangle \pm |V\rangle}{\sqrt{2}}$ as two of their states, then vary the azimuthal angle of the other two states sent $\frac{|H\rangle \pm e^{i\phi}|V\rangle}{\sqrt{2}}$, and observe how the key rate changes. By symmetry, we need only vary $\phi \in [0, \pi]$.

In the top plot of Fig. 10, we plot the results assuming a single-photon source and a distance of 10 km. As expected, the key rate is independent of ϕ when there is no leakage light. In the presence of leakage light, $\phi = \pi/2$ still remains as the optimum test state to send, but the key rate drops off away from that point, most dramatically for Model 1. Interestingly, there is even a region for which Model 2 outperforms Model 3. To explain this, we can go to the Gram matrix formed by the initial states which form the constraints on the RHS of Eq. (2). If we calculate the trace distance between the Gram matrix of Model 2 and the Gram matrix created by the ideal

qubit states $\{\frac{|H\rangle \pm |V\rangle}{\sqrt{2}}, \frac{|H\rangle \pm e^{i\phi}|V\rangle}{\sqrt{2}}\}$ as a function of ϕ , we find that it is symmetric about $\phi = \pi/2$; however, doing the same for the Gram matrix of Model 3, we find that the trace distance is not symmetric about that point due to the time-dependent nature of the underlying states and the way the inner product is calculated in Eq. (17). As ϕ increases, the Gram matrix of Model 3 eventually becomes a further distance from ideal than the Gram matrix of Model 2 for $\phi \gtrsim 0.8\pi$, so it is conceivable the key rate for Model 3 can perform worse in that region. Of course, the key rate depends on much more than just this trace distance, since the angle also changes the constraints provided by the detection statistics, but this provides some intuition as to why Model 2 can outperform Model 3 in certain regimes.

In bottom of Fig. 10, we plot the key rates assuming a decoy state method and a distance of 50 km. Here even the case of zero leakage light has some sensitivity to the angle of the test state. $\phi = \pi/2$ is still the optimal test state across all models. Like before, there is a limited range of ϕ that provides a positive key rate in the presence of leakage light, with the range being narrowest for Model 1. We observed for both types of sources that the range of ϕ that yields positive key rate narrows as the channel distance is increased; this means that source preparation flaws, especially in the test state, become a greater problem at further distances, unlike in the case of no leakage light where there is greater stability of the key rate with respect to ϕ .

The main point of these simulations is to demonstrate that while the choice of test state is not so important when the source is ideal without side channels, in the presence of leakage light, we must be careful to choose a test state that provides both good constraints on the encoded mode and on the leakage mode. While $\phi = \pi/2$ seemed to be the best choice for these models—coinciding with the BB84 states—we also observed cases when other values of ϕ produced the maximum key rate at a given distance. In a typical protocol, Alice and Bob simply choose the BB84 states and optimize the decoy state intensities as a function of distance; here, we see that in the presence of leakage light, there is additional benefit to optimizing over the polarization of the test states sent.

-
- [1] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani *et al.*, *Adv. Opt. Photon.* **12**, 1012 (2020).
- [2] H.-K. Lo, M. Curty, and B. Qi, *Phys. Rev. Lett.* **108**, 130503 (2012).
- [3] K. Tamaki, M. Curty, G. Kato, H.-K. Lo, and K. Azuma, *Phys. Rev. A* **90**, 052314 (2014).
- [4] Z. Tang, K. Wei, O. Bedroja, L. Qian, and H.-K. Lo, *Phys. Rev. A* **93**, 042308 (2016).
- [5] J. E. Bourassa, I. W. Primaatmaja, C. C. W. Lim, and H.-K. Lo, *Phys. Rev. A* **102**, 062607 (2020).
- [6] 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).
- [7] K. Tamaki, M. Curty, and M. Lucamarini, *New J. Phys.* **18**, 065008 (2016).
- [8] W. Wang, K. Tamaki, and M. Curty, *Sci. Rep.* **11**, 1 (2021).
- [9] M. Pereira, G. Kato, A. Mizutani, M. Curty, and K. Tamaki, *Sci. Adv.* **6**, eaaz4487 (2020).
- [10] A. Navarrete, M. Pereira, M. Curty, and K. Tamaki, *Phys. Rev. Appl.* **15**, 034072 (2021).
- [11] M. Pereira, M. Curty, and K. Tamaki, *npj Quantum Inf.* **5**, 62 (2019).
- [12] I. W. Primaatmaja, E. Lavie, K. T. Goh, C. Wang, and C. C. W. Lim, *Phys. Rev. A* **99**, 062332 (2019).
- [13] I. Lucio-Martinez, P. Chan, X. Mo, S. Hosier, and W. Tittel, *New J. Phys.* **11**, 095001 (2009).
- [14] Z. Tang, Z. Liao, F. Xu, B. Qi, L. Qian, and H.-K. Lo, *Phys. Rev. Lett.* **112**, 190503 (2014).
- [15] J. Wang, X. Qin, Y. Jiang, X. Wang, L. Chen, F. Zhao, Z. Wei, and Z. Zhang, *Opt. Express* **24**, 8302 (2016).
- [16] C. Li, M. Curty, F. Xu, O. Bedroja, and H.-K. Lo, *Phys. Rev. A* **98**, 042324 (2018).

- [17] E. Moschandreou, B. J. Rollick, B. Qi, and G. Siopsis, *Phys. Rev. A* **103**, 032614 (2021).
- [18] L. C. Comandar, M. Lucamarini, B. Fröhlich, J. F. Dynes, A. W. Sharpe, S. W.-B. Tam, Z. L. Yuan, R. V. Penty, and A. J. Shields, *Nat. Photon.* **10**, 312 (2016).
- [19] A. Duplinskiy and D. Sych, *Phys. Rev. A* **104**, 012601 (2021).
- [20] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, *Phys. Rev. A* **73**, 022320 (2006).
- [21] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. L. Yuan, and A. J. Shields, *Phys. Rev. X* **5**, 031030 (2015).
- [22] G. Zhang, I. W. Primaatmaja, J. Y. Haw, X. Gong, C. Wang, and C. C. W. Lim, *PRX Quantum* **2**, 030304 (2021).
- [23] In fact, the passive side channel we study in this paper is introduced by active optical elements controlled by Alice and Bob.
- [24] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, *Rev. Mod. Phys.* **92**, 025002 (2020).
- [25] C.-H. F. Fung, and H.-K. Lo, *Phys. Rev. A* **74**, 042342 (2006).
- [26] C. H. Bennett and G. Brassard, *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (Bangalore, India, 1984), pp. 175–179.
- [27] P. J. Coles, E. M. Metodiev, and N. Lütkenhaus, *Nat. Commun.* **7**, 11712 (2016).
- [28] A. Winick, N. Lütkenhaus, and P. J. Coles, *Quantum* **2**, 77 (2018).
- [29] N. T. Islam, C. C. W. Lim, C. Cahall, J. Kim, and D. J. Gauthier, *Phys. Rev. A* **97**, 042347 (2018).
- [30] Y. Wang, I. W. Primaatmaja, E. Lavie, A. Varvitsiotis, and C. C. W. Lim, *npj Quantum Inf.* **5**, 17 (2019).
- [31] N. T. Islam, C. C. W. Lim, C. Cahall, B. Qi, J. Kim, and D. J. Gauthier, *Quantum Sci. Technol.* **4**, 035008 (2019).
- [32] E. Y.-Z. Tan, R. Schwonnek, K. T. Goh, I. W. Primaatmaja, and C. C.-W. Lim, *Quantum Inf.* **7**, 158 (2021).
- [33] I. George, J. Lin, and N. Lütkenhaus, *Phys. Rev. Res.* **3**, 013274 (2021).
- [34] S. Diamond and S. Boyd, *J. Mach. Learn. Res.* **17**, 1 (2016).
- [35] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, *J. Control Decis.* **5**, 42 (2018).
- [36] P. W. Shor and J. Preskill, *Phys. Rev. Lett.* **85**, 441 (2000).
- [37] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, *Quantum Inf. Comput.* **4**, 136 (2003).
- [38] I. Devetak and A. Winter, *Proc. R. Soc. Lond. A* **461**, 207 (2005).
- [39] C. Zhou, G. Wu, X. Chen, and H. Zeng, *Appl. Phys. Lett.* **83**, 1692 (2003).
- [40] X. Ma, C.-H. F. Fung, and M. Razavi, *Phys. Rev. A* **86**, 052305 (2012).
- [41] X. Zhong, J. Hu, M. Curty, L. Qian, and H.-K. Lo, *Phys. Rev. Lett.* **123**, 100506 (2019).
- [42] H.-K. Lo, *Quantum Inf. Comput.* **1**, 81 (2001).
- [43] H.-K. Lo, X. Ma, and K. Chen, *Phys. Rev. Lett.* **94**, 230504 (2005).
- [44] C. M. Gee, G. D. Thurmond, and H. W. Yen, *Appl. Opt.* **22**, 2034 (1983).