Practical Quantum Key Distribution That is Secure Against Side Channels

Álvaro Navarrete,[1,*] Margarida Pereira,[1] Marcos Curty,[1] and Kiyoshi Tamaki[2]

[1] *EI Telecomunicación, Department of Signal Theory and Communications, University of Vigo, Vigo E-36310, Spain*

[2] *Faculty of Engineering, University of Toyama, Gofuku 3190, Toyama 930-8555, Japan*

(Received 8 July 2020; revised 15 November 2020; accepted 1 March 2021; published 24 March 2021)

There is a large gap between theory and practice in quantum key distribution (QKD) because real devices do not satisfy the assumptions required by the security proofs. Here, we close this gap by introducing a simple and practical measurement-device-independent-QKD type of protocol, based on the transmission of coherent light, for which we prove its security against any possible imperfection and/or side channel from the quantum communication part of the QKD devices. Our approach only requires to experimentally characterize an upper bound of one single parameter for each of the pulses sent, which describes the quality of the source. Moreover, unlike device-independent (DI) QKD, it can accommodate information leakage from the users' laboratories, which is essential to guarantee the security of QKD implementations. In this sense, its security goes beyond that provided by DI QKD, yet it delivers a secret key rate that is various orders of magnitude greater than that of DI QKD.

DOI: 10.1103/PhysRevApplied.15.034072

I. INTRODUCTION

Recent years have witnessed tremendous progress in the field of quantum key distribution (QKD) [1–3], which includes the realization of long-distance fiber-based implementations [4–6], satellite links [7–9], and the deployment of QKD networks [10–12]. Despite these groundbreaking results, however, the security of QKD implementations has not been fully established yet, due to the difficulty of real devices to satisfy the assumptions required by the security proofs.

To bridge this pressing gap between theory and practice in QKD, the ultimate solution is arguably device-independent (DI) QKD [13–15], because it can guarantee security without characterizing the internal functioning of the QKD apparatuses. This is achieved by using a loophole-free Bell test [16–18] to verify that the two QKD users (called Alice and Bob) share nonlocal correlations. Importantly, such a test can be evaluated using only the measurement statistics directly observed by Alice and Bob, and no characterization of their measurement devices is required. The main drawbacks of DI QKD are, however, its impracticability with current technology and its poor expected performance [19]. Also, security proofs of DI QKD typically assume that the users' devices do not leak any unwanted information to the eavesdropper (Eve), which might be very challenging to assure in practice. For instance, Eve might launch a Trojan-horse attack (THA) [20–22] to learn Alice and Bob's measurement outcomes each given time, or, say, device imperfections like, for example, the emission of backflash light by their detectors [23–25] might leak this information to Eve.

Here, we introduce an alternative approach to achieve implementation security in QKD. It is based on a much simpler experimental setup than that typically used in DI QKD. It requires, however, that Alice and Bob perform some minimal characterization of their setup. Precisely, they need to estimate an upper bound on a single experimental parameter for each of the pulses sent, being these parameters related to the quality of their sources. In doing so, we show that, in contrast to DI QKD, our approach guarantees security against *any* possible side channel from the quantum communication part of the QKD devices, including any unwanted information leakage about Alice's and Bob's internal settings to Eve. Moreover, our scheme can be implemented with low detection efficiency detectors, yet can deliver a secret key rate that is various orders of magnitude greater than that of DI QKD.

The key idea is rather simple. To protect against detection side channels, which are arguably the Achilles' heel of QKD [26,27], we use a measurement-device-independent (MDI) QKD type of scheme [28]. Indeed, the schematic of the protocol that we introduce below actually resembles that of a MDI-QKD variant called twin-field QKD [29–35]. We remark already here, however, that its secret key rate scales linearly with the channel transmittance, like the original MDI-QKD scheme. Note that the secret key rate of twin-field QKD scales with the square root of the channel

[*] anavarrete@com.uvigo.es

transmittance. On the other hand, to protect against source side channels, we combine recent techniques that can efficiently handle all types of device imperfections at the source [36,37]. They require a minimum characterization of Alice's and Bob's emitted signals, in order to incorporate this information into the security proof and guarantee security.

We note that the problem of device imperfections and/or side channels at the source has been addressed in various recent works. For instance, state preparation flaws (SPFs) can be efficiently incorporated into the security analysis by means of the loss-tolerant protocol [5,36,38,39]. Also, discrete phase randomization has been addressed in Ref. [40]. Moreover, techniques to investigate the problem of information leakage about Alice's and Bob's internal settings (due to, say, a THA) have been introduced in [36,41,42]. More recently, methods to analyze the effect of classical pulse correlations in high-speed QKD have been presented in Refs. [37,43]. However, none of these works has simultaneously considered all possible side channels in a practical QKD implementation, and thus they cannot guarantee implementation security. In this paper, we accommodate all of them.

II. PROTOCOL DESCRIPTION

For simplicity, in the protocol description we assume the ideal scenario where there are no side channels and all the prepared states are perfect. The presence of side channels or SPFs is discussed afterwards. That is, the description below represents an idealized scenario, and, in practice, Alice and Bob do not necessarily have to generate the states assumed here. Moreover, we consider the symmetric situation where the set of transmitted states and their *a priori* probabilities are equal for Alice and Bob. Also, we assume that the untrusted node Charles, which is required in MDI-QKD schemes to perform the measurements, is located in the middle between them. We remark, however, that the generalization to the asymmetric scenario is straightforward [44–49]. The setup is shown in Fig. 1.

1. Alice (Bob) sends the state $|\nu\rangle_a$ ($|\omega\rangle_b$) to the untrusted node Charles with probability p_ν (p_ω), where $\nu, \omega \in \mathcal{T} := \{\alpha, -\alpha, \text{vac}\}$. In particular, the key state $|\alpha\rangle$ ($|-\alpha\rangle$) is a coherent state with amplitude α ($-\alpha$) and it is associated with the bit value 0 (1), while the vacuum state $|\text{vac}\rangle$ is used for parameter estimation.

2. If Charles is honest, he causes the incoming pulses to interfere using a 50:50 beam splitter followed by two threshold detectors, D_c and D_d, which are associated with constructive and destructive interference, respectively. If his measurement succeeds, which means that only one of his detectors clicks, Charles announces the measurement outcome $\Omega \in \{\Omega_c, \Omega_d\}$, where Ω_c (Ω_d) corresponds to a click event only in detector D_c (D_d). Otherwise, he announces the failure event. Besides, if Ω_d is announced, Bob flips his bit value.

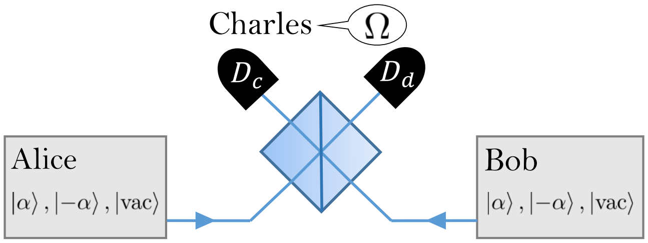
Charles
Ω
D_c
D_d
Alice
$|\alpha\rangle, |-\alpha\rangle, |\text{vac}\rangle$
Bob
$|\alpha\rangle, |-\alpha\rangle, |\text{vac}\rangle$

FIG. 1. Graphical illustration of the protocol. In each round, Alice and Bob each randomly select one state from the set $\{|\alpha\rangle, |-\alpha\rangle, |\text{vac}\rangle\}$ and send it to Charles, who causes the incoming signals to interfere using a 50:50 beam splitter followed by two threshold detectors, D_c and D_d. Here, $|\alpha\rangle$ and $|-\alpha\rangle$ denote coherent states and $|\text{vac}\rangle$ is the vacuum state.

3. The previous two steps are repeated N times. Next, Alice and Bob reveal their state choices for all the rounds in which at least one of them sent the vacuum state. The bits associated with the remaining rounds declared as successful by Charles constitute their sifted key.

4. Alice and Bob announce part of their sifted key and they estimate both the bit and the phase error rates. Finally, they perform error correction and privacy amplification to obtain, with high probability, a secret key.

III. SIDE CHANNELS

As already mentioned, being a MDI-QKD [28] type of protocol, the scheme above is immune against *all* detection side-channel attacks, so below we focus only on the potential side channels at the transmitters. We begin by explaining how we describe the emitted states, and then we move on to the security proof.

For each particular round of the protocol, if Alice and Bob select, say, the settings ν and ω, respectively, the joint state of their transmitted systems a and b, and Eve's system E, can always be written as

$$|\Psi_{\nu,\omega}\rangle_T = \sqrt{1-\epsilon_{\nu,\omega}}\,|\phi_{\nu,\omega}\rangle_T + \sqrt{\epsilon_{\nu,\omega}}\,|\phi^{\perp}_{\nu,\omega}\rangle_T, \qquad (1)$$

where $T := abE$, $\epsilon_{\nu,\omega} \in [0,1]$, $|\phi_{\nu,\omega}\rangle_T := |\nu\rangle_a |\omega\rangle_b |\tau\rangle_E$ with $|\tau\rangle_E$ being a state that does not contain any information about Alice's and Bob's setting choices for the current round, and $|\phi^{\perp}_{\nu,\omega}\rangle_T$ is a state orthogonal to $|\phi_{\nu,\omega}\rangle_T$. Importantly, as we show below, Eq. (1) represents the most general description of the transmitted states, which means that any potential SPF or information leakage about the internal settings of Alice and Bob can be characterized with this equation. This includes active information leakage due to, say, a THA [20,22,41,42,50], passive information leakage due to device imperfections, or both of them simultaneously. Also, it includes classical pulse correlations, since they can be treated as passive information leakage [37], as well as coherent attacks. To see this latter fact, one can consider the purification of all systems held by Alice, Bob, and Eve during the protocol. Moreover, we allow all systems held by Eve to jointly interact with all the optical pulses that have been emitted by Alice and Bob. Also, we introduce some ancilla systems $A\bar{A}$ and $B\bar{B}$ for Alice and Bob, respectively, that contain all their setting information in an entanglement-based picture of the protocol. Here we use the notation x ($\bar{x}$) to encapsulate the systems belonging to the particular round (all rounds already transmitted except the particular round) that is being considered. In doing so, the entire global system in the considered round comprises the systems $ABab\bar{A}\bar{B}\bar{a}\bar{b}E$. Now, if Alice and Bob perform projective measurements on their ancillas A and B to obtain their setting information for that particular round (note that systems $\bar{A}$ and $\bar{B}$ could have been already measured), it is straightforward to show that the resulting state for that round can be written in the form of Eq. (1) by simply redefining the joint system $\bar{A}\bar{B}\bar{a}\bar{b}E$ as E. In what follows, we use E to refer to this joint system $\bar{A}\bar{B}\bar{a}\bar{b}E$.

To explicitly show that Eq. (1) is indeed the most general description of the transmitted states, let $|\tilde{\varphi}_{\nu,\omega}\rangle_E$ be an unnormalized state such that $|\tilde{\varphi}_{\nu,\omega}\rangle_E = {}_a\langle\nu|_b\langle\omega|\,|\Psi_{\nu,\omega}\rangle_T$. Note that Eq. (1) holds trivially with $\epsilon_{\nu,\omega} = 1$ if ${}_a\langle\nu|_b\langle\omega|\,|\Psi_{\nu,\omega}\rangle_T = 0$. Then, $|\Psi_{\nu,\omega}\rangle_T$ can always be written in the form

$$|\Psi_{\nu,\omega}\rangle_T = |\nu\rangle_a\,|\omega\rangle_b\,|\tilde{\varphi}_{\nu,\omega}\rangle_E + |\tilde{\chi}_{\nu,\omega}\rangle_T\,, \tag{2}$$

with $|\tilde{\chi}_{\nu,\omega}\rangle_T$ another unnormalized state such that ${}_a\langle\nu|_b\langle\omega|\,|\tilde{\chi}_{\nu,\omega}\rangle_T = 0$. Similarly, for some $\epsilon_{\nu,\omega} \in [0, 1]$, the unnormalized state $|\tilde{\varphi}_{\nu,\omega}\rangle_E$ can always be written as

$$|\tilde{\varphi}_{\nu,\omega}\rangle_E = \sqrt{1-\epsilon_{\nu,\omega}}\,|\tau\rangle_E + |\tilde{\tau}^{\perp}_{\nu,\omega}\rangle_E\,, \tag{3}$$

where $|\tau\rangle_E$ is a normalized state that does not depend on the internal settings of the transmitters and $|\tilde{\tau}^{\perp}_{\nu,\omega}\rangle_E$ is an unnormalized state orthogonal to $|\tau\rangle_E$. Finally, by combining Eqs. (2) and (3), one directly recovers Eq. (1) with $\sqrt{\epsilon_{\nu,\omega}}\,|\phi^{\perp}_{\nu,\omega}\rangle_T = |\tilde{\chi}_{\nu,\omega}\rangle_T + |\nu\rangle_a\,|\omega\rangle_b\,|\tilde{\tau}^{\perp}_{\nu,\omega}\rangle_E$.

Let us conclude this part by further illustrating the meaning of Eq. (1) with a simple example. For instance, suppose that there is a THA where $|\Psi_{\nu,\omega}\rangle_T = |\nu\rangle_a\,|\omega\rangle_b\,|\Lambda_{\nu,\omega}\rangle_E$, with $|\Lambda_{\nu,\omega}\rangle_E$ the state of the back-reflected light that carries information about the transmitters' settings. We note, however, that $|\Lambda_{\nu,\omega}\rangle_E$ could also represent any other type of side channel from the quantum communication part of the QKD devices, such as, for instance, passive electromagnetic or acoustic radiation. The state $|\Lambda_{\nu,\omega}\rangle_E$ can always be written as a superposition of the vacuum state and a state $|\Lambda'_{\nu,\omega}\rangle_E$ that contains no vacuum component, i.e., $|\Psi_{\nu,\omega}\rangle_T = \lambda\,|\nu\rangle_a\,|\omega\rangle_b\,|\text{vac}\rangle_E + \sqrt{1-\lambda^2}\,|\nu\rangle_a\,|\omega\rangle_b\,|\Lambda'_{\nu,\omega}\rangle_E$. This is so due to inevitable losses at the transmitters (e.g., produced by material absorption or due to the presence of optical isolators), which guarantee $\lambda > 0$. This latter equation is equivalent to Eq. (1) for $|\tau\rangle_E = |\text{vac}\rangle_E$.

IV. SECURITY PROOF

To prove the security of the protocol above, we assume that Alice and Bob know an upper bound on $\epsilon_{\nu,\omega}$ for each round, which is dependent on the previous setting choices, but no characterization is needed for the side-channel states $|\phi^{\perp}_{\nu,\omega}\rangle_T$ in Eq. (1). We remark, however, that any available information about the states $|\phi^{\perp}_{\nu,\omega}\rangle_T$ could be readily incorporated in the security proof described below. Also, we emphasize that the security proof is valid even if the states that Alice and Bob generate in the ideal scenario (i.e., without side channels) are not $|\nu\rangle_a$ and $|\omega\rangle_b$, or they are mixed states, due, for instance, to SPFs. In other words, $|\nu\rangle_a$ and $|\omega\rangle_b$ are adopted in Eq. (1) just as a reference for the state characterization in the experiment.

To calculate a lower bound on the secret key rate of the protocol, we first need to estimate the phase error rate e_{ph}, which is a key parameter in the complementarity argument [51]. For this, note that, from Eve's perspective, any of the rounds used for key generation in which both Alice and Bob send Charles key states (i.e., the state $|\alpha\rangle$ or $|-\alpha\rangle$ in the absence of side channels) is equivalently described by a fictitious scenario in which, instead, they first prepare the entangled state

$$|\Psi^{\text{vir}}\rangle_{ABT} = \frac{1}{2}\sum_{j,s=0,1} |j_z, s_z\rangle_{AB}\,|\Psi_{(-1)^j\alpha,(-1)^s\alpha}\rangle_T \tag{4}$$

with $\{|0_z\rangle\,, |1_z\rangle\}$ being the computational basis for the ancilla systems A and B, and subsequently they send the system T to Charles. This equivalence holds because measurements on ancilla systems A and B commute with those on system T. Note, moreover, that if Alice and Bob measure the ancilla systems of $|\Psi^{\text{vir}}\rangle_{ABT}$ in the computational basis, they randomly prepare the four possible joint key states $|\Psi_{(-1)^j\alpha,(-1)^s\alpha}\rangle_T$ (now including the side channels) that are sent in the actual protocol. Here, and in what follows, we consider that $j, s \in \{0, 1\}$ when referring to the virtual states. Now, we can imagine a fictitious virtual scenario where Alice and Bob measure their ancillas A and B in the complementary basis $\{|0_x\rangle\,, |1_x\rangle\}$, with $|j_x\rangle = (1/\sqrt{2})[|0_z\rangle + (-1)^j\,|1_z\rangle]$. In this virtual scenario, the unnormalized reduced density operators of the transmitted states are given by

$$\bar{\sigma}^{\text{vir}}_{j,s} = \text{Tr}_{AB}[|j_x, s_x\rangle\langle j_x, s_x|_{AB} \otimes \mathbb{1}_T |\Psi^{\text{vir}}\rangle\langle\Psi^{\text{vir}}|_{ABT}], \tag{5}$$

where $\mathbb{1}_T$ is the identity operator acting on T. We call the states $\bar{\sigma}^{\text{vir}}_{j,s}$ the unnormalized virtual states, and we write their normalized form as $\sigma^{\text{vir}}_{j,s} \equiv |\Psi^{\text{vir}}_{j,s}\rangle\langle\Psi^{\text{vir}}_{j,s}|_T$.

The phase error rate is then defined as the bit error rate of the virtual scenario. In the protocol above, a phase error occurs when Alice and Bob measure either $|0_x, 0_x\rangle_{AB}$ or $|1_x, 1_x\rangle_{AB}$ and Charles announces a successful event (see Appendix A). This means that

$$e_{\text{ph}} = \frac{p_{0,0}^{\text{vir}} Y_{0,0}^{\text{vir}} + p_{1,1}^{\text{vir}} Y_{1,1}^{\text{vir}}}{\sum_{j,s} p_{j,s}^{\text{vir}} Y_{j,s}^{\text{vir}}}, \tag{6}$$

where $Y_{j,s}^{\text{vir}}$ is the conditional probability of a successful announcement by Charles given that Alice and Bob send $\sigma_{j,s}^{\text{vir}}$, and $p_{j,s}^{\text{vir}} = \text{Tr}\{\bar{\sigma}_{j,s}^{\text{vir}}\}$. Note that, since Alice and Bob measure their ancillas in the complementary basis, the bit flip operation performed by Bob when Charles announces a result Ω_d has no effect in the virtual scenario. The term $\sum_{j,s} p_{j,s}^{\text{vir}} Y_{j,s}^{\text{vir}} =: \gamma_{\text{obs}}$ in Eq. (6) is equal to the probability that Charles announces a successful event and both Alice and Bob send a key state. This quantity is directly observed in the actual experiment. Thus, to calculate e_{ph}, it is enough to estimate the phase error probability $p_{0,0}^{\text{vir}} Y_{0,0}^{\text{vir}} + p_{1,1}^{\text{vir}} Y_{1,1}^{\text{vir}} =: \Gamma$.

For this, we use the reference technique recently introduced in Ref. [37]. Specifically, we first define, for each user, a set of qubit states $\{|\Phi_\alpha\rangle, |\Phi_{-\alpha}\rangle, |\Phi_{\text{vac}}\rangle\}$ called the reference states. We have freedom to select the reference states; however, for the security proof to go through, a lower bound on $|\langle\Phi_{\nu,\omega}||\Psi_{\nu,\omega}\rangle|$ for each possible combination of ν and ω is needed, with $|\Phi_{\nu,\omega}\rangle_T := |\Phi_\nu\rangle_a \otimes |\Phi_\omega\rangle_b \otimes |\tau\rangle_E$. That is, the joint reference states $|\Phi_{\nu,\omega}\rangle_T$ should be chosen similar to the original transmitted states $|\Psi_{\nu,\omega}\rangle_T$, which in practice is equivalent to saying that they should be similar to the states $|\phi_{\nu,\omega}\rangle_T$. We remark that the reference states are never prepared nor sent in the actual protocol, but they are used here just as a mathematical tool to find an upper bound on Γ that depends on the statistics of the original transmitted states. In what follows, we omit the mode subscripts for readability whenever it is clear from the context.

A natural choice for the set of reference states is given by $\{|\alpha\rangle, |-\alpha\rangle, |\text{vac}'\rangle\}$, where the state $|\text{vac}'\rangle$ is the projection of $|\text{vac}\rangle$ onto the qubit space spanned by $\{|\alpha\rangle, |-\alpha\rangle\}$. For this, let the orthonormal basis $\{|0_o\rangle, |1_o\rangle, |2_o\rangle\}$ satisfy

$$\begin{aligned}
|\alpha\rangle &= |0_o\rangle\,, \\
|-\alpha\rangle &= \langle\alpha|-\alpha\rangle\,|0_o\rangle + \sqrt{1-|\langle\alpha|-\alpha\rangle|^2}\,|1_o\rangle\,, \\
|\text{vac}\rangle &= \langle\alpha|\text{vac}\rangle\,|0_o\rangle + c_1\,|1_o\rangle + c_2\,|2_o\rangle\,,
\end{aligned}$$

where the coefficients c_1 and c_2 fulfil $\langle-\alpha|\text{vac}\rangle = \langle-\alpha|\alpha\rangle\langle\alpha|\text{vac}\rangle + c_1\sqrt{1-|\langle\alpha|-\alpha\rangle|^2}$ and $|\langle\alpha|\text{vac}\rangle|^2 + |c_1|^2 + c_2^2 = 1$, and where, without loss of generality, we assume that c_2 is real. This means, in particular, that $|\text{vac}'\rangle = (1/\sqrt{\xi})[\langle\alpha|\text{vac}\rangle\,|0_o\rangle + c_1\,|1_o\rangle]$, with $\xi = |\langle\alpha|\text{vac}\rangle|^2 + |c_1|^2$.

From the definitions of $|\Psi^{\text{vir}}\rangle$, $\sigma_{j,s}^{\text{vir}}$, $p_{j,s}^{\text{vir}}$, and $Y_{j,s}^{\text{vir}}$, one can define analogous states and probabilities $|\Phi^{\text{vir}}\rangle$, $\sigma_{j,s}^{\text{vir}|\text{ref}}$, $p_{j,s}^{\text{vir}|\text{ref}}$, and $Y_{j,s}^{\text{vir}|\text{ref}}$ for the reference states above by simply substituting the actual states $|\Psi_{\nu,\omega}\rangle$ with the reference states $|\Phi_{\nu,\omega}\rangle$ where needed in their definitions [37]. In particular, the yields $Y_{j,s}^{\text{vir}|\text{ref}}$ are defined as

$$Y_{j,s}^{\text{vir}|\text{ref}} = \text{Tr}\big[\hat{\mathcal{D}}\sigma_{j,s}^{\text{vir}|\text{ref}}\big], \tag{7}$$

where $\hat{\mathcal{D}}$ is the positive operator-valued measure element associated with Charles' successful announcement. Now, to estimate Γ, one can define an analogous quantity for the reference states, namely $\Gamma_{\text{ref}} := p_{0,0}^{\text{vir}|\text{ref}} Y_{0,0}^{\text{vir}|\text{ref}} + p_{1,1}^{\text{vir}|\text{ref}} Y_{1,1}^{\text{vir}|\text{ref}}$, and then quantify the maximum possible deviation in the measurement statistics between the reference and actual scenarios. For this, we conveniently define the operator $\hat{\mathcal{D}}_{\text{ph}} = (|0_x, 0_x\rangle\langle 0_x, 0_x| + |1_x, 1_x\rangle\langle 1_x, 1_x|) \otimes \hat{\mathcal{D}}$ and then we use the fact that, for any operator $0 \preceq \hat{\mathcal{O}} \preceq \mathbb{1}$, and normalized pure states $|A\rangle$ and $|R\rangle$, the following inequality is satisfied [37]:

$$\delta \leq \sqrt{Y_A Y_R} + \sqrt{(1-Y_A)(1-Y_R)}. \tag{8}$$

Here $\delta = |\langle A|R\rangle|$, $Y_A = \langle A|\,\hat{\mathcal{O}}\,|A\rangle$, and $Y_R = \langle R|\,\hat{\mathcal{O}}\,|R\rangle$. From Eq. (8) one can derive the functions

$$G_+(Y_R, \delta) = \begin{cases} g_+(Y_R, \delta), & Y_R < \delta^2, \\ 1, & \text{otherwise}, \end{cases} \tag{9}$$

$$G_-(Y_R, \delta) = \begin{cases} g_-(Y_R, \delta), & Y_R > 1-\delta^2, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

such that $G_-(Y_R, \delta) \leq Y_A \leq G_+(Y_R, \delta)$, where $g_\pm(Y, \delta) = Y + (1-\delta^2)(1-2Y) \pm 2\delta\sqrt{(1-\delta^2)Y(1-Y)}$. Furthermore, given $Y^U \geq Y$ and $0 \leq \delta^L \leq \delta$, it holds that $G_+(Y^U, \delta^L) \geq G_+(Y, \delta)$. Then, by noting that $\Gamma = \langle\Psi^{\text{vir}}|\,\hat{\mathcal{D}}_{\text{ph}}\,|\Psi^{\text{vir}}\rangle$ and $\Gamma_{\text{ref}} = \langle\Phi^{\text{vir}}|\,\hat{\mathcal{D}}_{\text{ph}}\,|\Phi^{\text{vir}}\rangle$, an upper bound on Γ can be simply obtained as

$$\Gamma \leq G_+(\Gamma_{\text{ref}}, \delta_{\text{vir}}) \leq G_+(\Gamma_{\text{ref}}^U, \delta_{\text{vir}}^L) =: \Gamma^U, \tag{11}$$

where Γ_{ref}^U is an upper bound on Γ_{ref} and $\delta_{\text{vir}}^L = (1/4)\sum_{j,s=0,1}\sqrt{1-\epsilon_{(-1)^j\alpha,(-1)^s\alpha}}$ is a lower bound on $\delta_{\text{vir}} := |\langle\Phi^{\text{vir}}|\Psi^{\text{vir}}\rangle|$. In particular, an upper bound Γ_{ref}^U can be calculated from all the observed statistics $Y_{\nu,\omega}$ of the actual states, for $\nu, \omega \in \{\alpha, -\alpha, \text{vac}\}$, and from the square root fidelities between the actual and reference states $|\langle\Psi_{\nu,\omega}|\Phi_{\nu,\omega}\rangle|$ (see Appendix B for a particular expression).

Importantly, it can be shown that Γ^U can be written as a concave function of the observed statistics $Y_{\nu,\omega} := \langle\Psi_{\nu,\omega}|\,\hat{\mathcal{D}}\,|\Psi_{\nu,\omega}\rangle$ and, therefore, the security of the protocol can be easily extended against coherent attacks. We refer the reader to Appendix C for further details. Also, note that in this security proof we consider that Alice and Bob measure the complementary observable of all their ancilla systems in the entanglement-based virtual protocol. This means that, for each round of the virtual protocol, the previous setting information is not available and, therefore, the state of the current round can be regarded as a classical mixture of states, each of which is dependent on the previous settings. Our security proof is valid because we consider the worst pure state, i.e., the state with an upper bound on $\epsilon_{\nu,\omega}$ that results in the worst probability for the phase error [recall that $G_+(Y,\delta) \leq G_+(Y,\delta^{\mathrm{L}})$], and importantly this upper bound is independent of the previous settings. Moreover, since we do not exploit the previous setting information in the security proof, the data processing in the actual protocol should not depend on it.

Finally, given an upper bound $e_{\mathrm{ph}}^{U} = \Gamma^{U}/\gamma_{\mathrm{obs}}$ on e_{ph}, the asymptotic secret key rate can be written as

$$R \geq Q[1 - h(e_{\mathrm{ph}}^{U}) - f_e h(e_{\mathrm{bit}})], \tag{12}$$

where e_{bit} is the bit error rate, f_e is the error correction efficiency, and Q is the probability that both Alice and Bob select a key state and Charles announces a successful event.

V. EVALUATION

In Fig. 2 we show the secret key rate of the protocol in the presence of side channels (solid lines). For simplicity, here we set $\epsilon_{\nu,\omega} = \epsilon$ for all $\nu, \omega \in \mathcal{T}$, and we optimize the parameter α for each value of the overall system loss. In our simulations, we model system loss with a beam splitter and, also, for simplicity, we disregard any misalignment effect in the channel. Note that, in practice, Alice and Bob need to maintain phase stability between their signals, which means that phase locking of remote laser pulses is required to properly perform the joint measurements at Charles' station [6,52–55]. In addition, we set the dark-count probability of Charles' detectors to $p_d = 10^{-8}$ to match some recent experiments [52]. Further details about the channel model and the optimal values for α can be found in Appendix D.

As expected, the performance of the protocol decreases when ϵ increases. Also, Fig. 2 shows that, for the channel model considered, a positive secret key rate is possible even when $\epsilon = 10^{-5}$. Note that ϵ characterizes, for each state $|\Psi_{\nu,\omega}\rangle$, the information leakage of both users. For instance, when $\epsilon = 10^{-6}$, our simulation results suggest that Alice and Bob could generate a secret key over about 14 dB of overall system loss, which corresponds to a transmission distance of about 50 km when considering threshold detectors with 44% of detection efficiency [52] and standard optical fibres with loss coefficient 0.2 dB/km. Moreover, we remark that this is achieved without requiring the use of the decoy-state technique [57–59] or the use of phase randomized coherent pulses, which could open additional side channels that Eve might exploit [41,42,60–62]. Furthermore, we note that Charles' station is simpler than those of other MDI-QKD type schemes, as it only requires two detectors (rather than four) to distinguish two Bell states [29,63], similar to twin-field QKD.

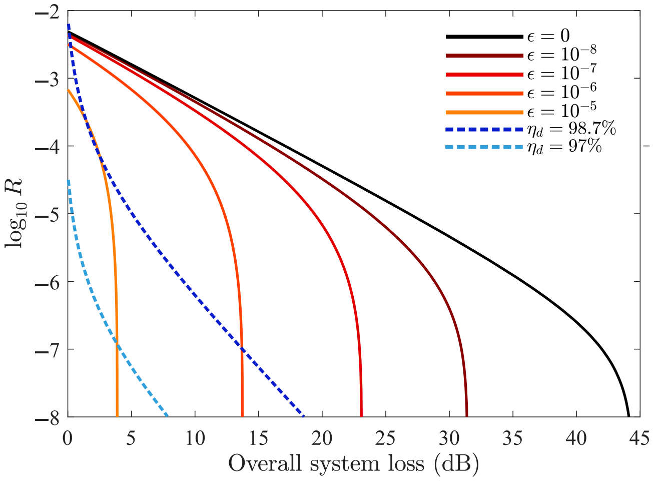

FIG. 2. Secret key rate R as a function of the overall system loss (in decibels) between Alice and Bob for different values of the parameter ϵ (solid lines). For simplicity, we consider the symmetric scenario where Charles is located in the middle between Alice and Bob. The value of α has been optimized for each system loss value. For comparison, the figure also includes the secret key rate of a CHSH-based DI-QKD protocol with parametric down-conversion sources, qubit amplifiers [56], and photon-number-resolving detectors for two different values of their detection efficiency η_d (dashed lines) [19]. Given that ϵ is sufficiently small, the resulting secret key rate of our protocol can be various orders of magnitude greater than that of DI QKD.

For comparison purposes, Fig. 2 also includes the asymptotic secret key rate of a Clauser-Horne-Shimony-Holt- (CHSH) [17] based DI-QKD protocol that uses parametric down-conversion sources and qubit amplifiers (dashed lines). Here, we consider the use of qubit amplifiers because, otherwise, the achievable distance and rate of DI QKD would be too poor to have any practical relevance. Precisely, we assume that Alice and Bob use an entanglement swapping relay [19,56]. This relay includes a Bell state measurement (BSM) in which independent light pulses interfere at a 50:50 beam splitter before they are projected into orthogonal polarization states by means of two polarizing beam splitters and four detectors. Moreover, to further enhance the performance of DI QKD, we consider that all photodetectors both in Alice's and Bob's labs, as well as within the BSM, are photon-number-resolving (PNR) detectors. We remark that this contrasts with our protocol, in which we assume the use of threshold detectors. Precisely, we consider two particular values for the detection efficiency of all PNR detectors, $\eta_d = 98.7\%$ and $\eta_d = 97\%$, and disregard the effect of coupling losses. For lower values of the detection efficiency, the performance of DI QKD decreases very rapidly and may result in no secret ket rate. Importantly, Fig. 2 shows that, even in this optimistic scenario, the secret key rate of DI QKD can be various orders of magnitude lower than that of our protocol above. What is more, as already mentioned in the Introduction, DI QKD assumes that Alice's and Bob's devices are perfectly isolated from the channel and there is no unwanted information leakage to Eve. In practice, however, this might be very challenging to achieve experimentally, and one expects that the presence of information leakage would decrease the resulting secret key rate even further, though it is unclear how to take information leakage into account in a security analysis of a DI-QKD protocol where the devices are uncharacterized.

As a side remark, we also note that if ϵ is sufficiently small, the tolerance of our protocol against the system loss becomes comparable to some MDI-QKD protocols that assume that the transmitted states are characterized precisely [39,64]. Also, its key rate is greater than that of the standard MDI-QKD scheme assuming leaky sources [65], even though this latter work assumes that there are no SPFs, Alice and Bob apply perfect phase randomization, and this phase information is not leaked to Eve.

To conclude this part, let us emphasize that the simulations in Fig. 2 assume that Alice and Bob can emit perfect vacuum signals when $\epsilon = 0$ and we consider side channels attached to the perfect vacuum states only for simplicity. In practice, however, due to the finite extinction ratio of intensity modulators, it might be difficult for Alice and Bob to generate a perfect vacuum state. Importantly, we note that very similar results as those illustrated in Fig. 2 can be obtained if we replace the vacuum signals with sufficiently weak coherent states (see Appendix E for further details).

VI. DISCUSSION

Our analysis demonstrates that, in principle, it is possible to incorporate all potential side channels of a QKD setup into a security proof. This is remarkable, and goes beyond the current security framework provided by DI QKD, where one typically assumes the absence of information leakage from the users' devices. Indeed, our result also has implications in cryptography in general, where a fundamental open question is whether or not it is possible (even theoretically) to achieve the Holy Grail of cryptography—i.e., information-theoretic security—with a practical setup. We have simplified this fundamental question to the problem of upper bounding certain parameters $\epsilon_{\nu,\omega}$ that concentrate all the potential side-channel information.

We admit however that obtaining an upper bound $\epsilon_{\nu,\omega}$ might still be a quite difficult task in practice. This is so because the Hilbert space associated with certain side channels could, in general, be infinite dimensional. In this regard, let us emphasize that this difficulty seems to be unavoidable, and comes from the generality of the problem. That is, it is not exclusive to the QKD protocol introduced in this manuscript or of its security proof. Indeed, if this important problem did not appear in previous security analyses of QKD, it is simply because it has been disregarded so far.

To present a precise method to characterize $\epsilon_{\nu,\omega}$ is clearly important for the field of cryptography, but it is beyond the scope of this work. Still, we firmly believe that upper bounding $\epsilon_{\nu,\omega}$ should definitely be far easier than the alternative of fully characterizing the side-channel states. In this sense, our protocol and security analysis represent an important step towards achieving full implementation security, and it is now the time for theorists to work together with experimentalists to design efficient and practical procedures to upper bound these parameters.

VII. CONCLUSIONS

We have introduced a MDI-QKD type of protocol and we have proven its security against any possible side channel from the quantum communication part of the QKD devices, thus closing the large gap between theoretical security and implementation security in QKD. Our protocol represents an alternative solution to that of DI QKD to achieve implementation security. Moreover, unlike DI QKD, it can accommodate information leakage from the users' laboratories in the security analysis, which is crucial to guarantee the security of the implementations. Also, our approach employs a much simpler experimental setup, and can offer a secret key rate that is orders of magnitude greater than that of DI QKD with practical signals. It requires, however, to experimentally estimate an upper bound on certain parameters that describe the quality of the sources. These parameters basically account for all source imperfections and side channels. That is, as long as such upper bounds are satisfied, the protocol is secure against any possible side channel from the quantum communication part of the QKD devices. In this regard, our approach offers a clear path for achieving secure QKD implementations with high performance.

ACKNOWLEDGMENTS

We thank Guillermo Currás-Lorenzo, Akihiro Mizutani, and Víctor Zapatero for valuable discussions. This work is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant No. 675662 (Project QCALL). M.C. also acknowledges support from the Spanish Ministry of Economy and Competitiveness (MINECO), and the Fondo Europeo de Desarrollo Regional (FEDER) through Grant No. TEC2017-88243-R. A.N. acknowledges support through a FPU fellowship (FPU 15/03668) from the Spanish Ministry of Science, Innovation and Universities (MCIU). K.T. acknowledges support from JSPS KAKENHI Grants No. JP18H05237 18H05237 and No. JST-CREST JPMJCR 1671.

All authors discussed the main idea and A.N. performed the analytical calculations and the numerical simulations. All authors analyzed the results and prepared the manuscript.

APPENDIX A: CHANNEL MODEL

Here we present the expected values for the quantities $Y_{\nu,\omega}$ used to estimate the phase error rate. For this, we model the loss from Alice (Bob) to Charles with a beam splitter of transmittance $\sqrt{\eta}$, i.e., the overall system loss is equal to $10\log(1/\eta)$. We further assume, for simplicity, that Charles' detectors have the same dark-count probability p_d, and we disregard the effect of phase misalignment introduced by the channel. In this scenario, it can be shown that the conditional probability that Charles observes a click in detector D_c but not in detector D_d given that Alice and Bob send him the states $|\nu\rangle$ and $|\omega\rangle$, respectively, is given by

$$Y_{\nu,\omega} = (1-p_d)^2 \exp\left[-\sqrt{\eta}\left(\frac{|\nu|^2+|\omega|^2}{2} - |\nu||\omega|\cos(\phi_A-\phi_B)\right)\right] \times \left\{1-\exp\left[-\sqrt{\eta}\left(\frac{|\nu|^2+|\omega|^2}{2} + |\nu||\omega|\cos(\phi_A-\phi_B)\right)\right]\right\} + p_d(1-p_d), \tag{A1}$$

where $\phi_A = \arg(\nu)$ and $\phi_B = \arg(\omega)$. The same probability given by Eq. (A1) is valid for the case where Charles observes destructive interference if one takes into account the bit flip at Bob's side (which is equivalent to flipping the phase of ω). On the other hand, the bit error rate is given by

$$e_{\text{bit}} = \frac{p_d}{2p_d + e^{2\sqrt{\eta}\alpha^2} - 1}. \tag{A2}$$

1. Phase error

Here we sketch how to decide the most convenient definition of a phase error in this protocol. For this, we assume the ideal scenario without side channels. This means that, in the entanglement-based picture, the state shared by Alice and Bob in the key rounds can be written as

$$|\Psi^{\text{vir}}\rangle = \tfrac{1}{2}[|0_z0_z\rangle_{AB}|\alpha,\alpha\rangle_{ab} + |0_z1_z\rangle_{AB}|\alpha,-\alpha\rangle_{ab} + |1_z0_z\rangle_{AB}|-\alpha,\alpha\rangle_{ab} + |1_z1_z\rangle_{AB}|-\alpha,-\alpha\rangle_{ab}]. \tag{A3}$$

The beam splitter at Charles acts on the input modes a and b as $\hat{a}^\dagger \to (1/\sqrt{2})[\hat{c}^\dagger + \hat{d}^\dagger]$ and $\hat{b}^\dagger \to (1/\sqrt{2})[\hat{c}^\dagger - \hat{d}^\dagger]$, with c and d the output modes corresponding to constructive and destructive interference, respectively, and where $\hat{m}^\dagger$ denotes the creation operator on mode m. Then, in an ideal scenario with no loss, the state after the beam splitter can be written as

$$|\Psi^{\text{vir}}\rangle = \tfrac{1}{2}[|0_z0_z\rangle_{AB}|\sqrt{2}\alpha\rangle_c + |1_z1_z\rangle_{AB}|-\sqrt{2}\alpha\rangle_c] + \tfrac{1}{2}[|0_z1_z\rangle_{AB}|\sqrt{2}\alpha\rangle_d + |1_z0_z\rangle_{AB}|-\sqrt{2}\alpha\rangle_d]. \tag{A4}$$

This means that the state associated with a click in D_c and no click in D_d is given by

$$|\Psi_c\rangle = \frac{e^{\alpha^2}}{\sqrt{1-e^{2\alpha^2}}}\sum_{\substack{n=1\\ n\ \text{odd}}}^{\infty}\left[\frac{(\sqrt{2}\alpha)^n}{n!}|n\rangle_c\right]\otimes\frac{1}{\sqrt{2}}(|0_z0_z\rangle_{AB} - |1_z1_z\rangle_{AB}) + \frac{e^{\alpha^2}}{\sqrt{1-e^{2\alpha^2}}}\sum_{\substack{n=2\\ n\ \text{even}}}^{\infty}\left[\frac{(\sqrt{2}\alpha)^n}{n!}|n\rangle_c\right]\otimes\frac{1}{\sqrt{2}}(|0_z0_z\rangle_{AB} + |1_z1_z\rangle_{AB}), \tag{A5}$$

where $|n\rangle_c$ is the Fock state with n photons on mode c. The previous state can be approximated, for small α, as

$$\begin{aligned}|\Psi_c\rangle &\approx \frac{e^{\alpha^2}\alpha}{\sqrt{1-e^{2\alpha^2}}}[|0_z0_z\rangle_{AB}-|1_z1_z\rangle_{AB}]\otimes|1\rangle_c\\ &= \frac{e^{\alpha^2}\alpha}{\sqrt{1-e^{2\alpha^2}}}[|0_x1_x\rangle_{AB}+|1_x0_x\rangle_{AB}]\otimes|1\rangle_c\,. \qquad \text{(A6)}\end{aligned}$$

Similarly, we can obtain exactly the same result for D_d if we take into account Bob's bit flip. This indicates that a phase error should be defined by Alice and Bob as observing identical outcomes (i.e., either $|0_x0_x\rangle$ or $|1_x1_x\rangle$) in the virtual scenario.

APPENDIX B: DERIVATION OF Γ^U_{ref}

Here we show how to obtain a simple upper bound on the quantity Γ_{ref}. For this, we first relate this quantity to the probabilities $Y^{\text{vir}|\text{ref}}_{j,s}$. We do so by rewriting the virtual states $\sigma^{\text{vir}|\text{ref}}_{j,s}$ as $\sigma^{\text{vir}|\text{ref}}_{j,s} = (1/4)\sum_{i,k} S^{j,s|\text{vir}}_{i,k}\sigma^a_i\otimes\sigma^b_k$, where σ^a_i and σ^b_k are the Pauli operators with $i,k\in\{\mathcal{I},X,Z\}$ and the terms $S^{j,s|\text{vir}}_{i,k}$ are the Bloch coefficients of the virtual states $\sigma^{\text{vir}|\text{ref}}_{j,s}$. Here the Pauli operator σ_Y is not necessary because none of the states $\sigma^{\text{vir}|\text{ref}}_{j,s}$ has complex components. Thus, Eq. (7) can be rewritten as

$$Y^{\text{vir}|\text{ref}}_{j,s} = \sum_{i,k} S^{j,s|\text{vir}}_{i,k} q_{i,k}, \qquad \text{(B1)}$$

where $q_{i,k} = (1/4)\text{Tr}[\hat{\mathcal{D}}\sigma^a_i\otimes\sigma^b_k]$. With this notation, one can conveniently write the matrix equation

$$\Gamma_{\text{ref}} = (\mathbf{P}^{\text{vir}})^{\text{T}}\mathbf{S}^{\text{vir}}\mathbf{q}, \qquad \text{(B2)}$$

where $(\mathbf{P}^{\text{vir}})^{\text{T}} = [p^{\text{vir}|\text{ref}}_{0,0}, p^{\text{vir}|\text{ref}}_{1,1}]$, $\mathbf{S}^{\text{vir}}$ is a 2 × 9 matrix containing the coefficients $S^{0,0|\text{vir}}_{i,k}$ ($S^{1,1|\text{vir}}_{i,k}$) in its first (second) row, and $\mathbf{q}$ is a column vector containing the quantities $q_{i,k}$. Moreover, and analogously to Eq. (B1), one can write

$$Y^{\text{ref}}_{\nu,\omega} = \sum_{i,k} S^{\nu,\omega}_{i,k} q_{i,k}, \qquad \text{(B3)}$$

where the $S^{\nu,\omega}_{i,k}$ denote the Bloch coefficients of the reference states $|\Phi_{\nu,\omega}\rangle$ and the quantities $Y^{\text{ref}}_{\nu,\omega}$ are their respective yields. From Eq. (B3), we find another matrix equation involving $\mathbf{q}$. It reads

$$\mathbf{Y}^{\text{ref}} = \mathbf{S}\mathbf{q}, \qquad \text{(B4)}$$

where $\mathbf{Y}^{\text{ref}}$ is a column vector containing the yields $Y^{\text{ref}}_{\nu,\omega}$ and $\mathbf{S}$ is a 9 × 9 matrix containing the Bloch coefficients of the reference states $|\Phi_{\nu,\omega}\rangle$ in its rows. Then, by combining Eqs. (B2) and (B4), one obtains

$$\Gamma_{\text{ref}} = (\mathbf{P}^{\text{vir}})^{\text{T}}\mathbf{S}^{\text{vir}}\mathbf{S}^{-1}\mathbf{Y}^{\text{ref}} = \mathbf{f}_{\text{obj}}\mathbf{Y}^{\text{ref}}, \qquad \text{(B5)}$$

where $\mathbf{f}_{\text{obj}} := (\mathbf{P}^{\text{vir}})^{\text{T}}\mathbf{S}^{\text{vir}}\mathbf{S}^{-1}$ is a row vector. Note that the matrix $\mathbf{S}$ is invertible because it can be written as the tensor product of two 3 × 3 invertible matrices.

Now, to obtain an upper bound on Γ_{ref}, we bound each term in Eq. (B5) separately. Specifically, we have

$$\begin{aligned}\Gamma_{\text{ref}} &= \mathbf{f}_{\text{obj}}\mathbf{Y}^{\text{ref}}\\ &= \sum_{\nu,\omega} f_{\nu,\omega}Y^{\text{ref}}_{\nu,\omega}\\ &\leq \sum_{\nu,\omega|f_{\nu,\omega}>0} f_{\nu,\omega}G_+(Y_{\nu,\omega},\delta^L_{\nu,\omega})\\ &\quad+\sum_{\nu,\omega|f_{\nu,\omega}<0} f_{\nu,\omega}G_-(Y_{\nu,\omega},\delta^L_{\nu,\omega})\\ &=: \Gamma^U_{\text{ref}}, \qquad \text{(B6)}\end{aligned}$$

where the coefficients $f_{\nu,\omega}$ are the elements of the vector $\mathbf{f}_{\text{obj}}$, the observed statistics $Y_{\nu,\omega} = \langle\Psi_{\nu,\omega}|\,\hat{\mathcal{D}}\,|\Psi_{\nu,\omega}\rangle$, and the terms $\delta^L_{\nu,\omega}$ are lower bounds on $\delta_{\nu,\omega} = |\langle\Psi_{\nu,\omega}|\Phi_{\nu,\omega}\rangle|$. To obtain particular expressions for the latter bounds, we first note that $|\langle\Psi_{\nu,\omega}|\Phi_{\nu,\omega}\rangle| = |\sqrt{1-\epsilon_{\nu,\omega}}\langle\phi_{\nu,\omega}|\Phi_{\nu,\omega}\rangle + \sqrt{\epsilon_{\nu,\omega}}\langle\phi^\perp_{\nu,\omega}|\Phi_{\nu,\omega}\rangle|$. Now, the reference states $|\Phi_{\nu,\omega}\rangle$ can always be written as

$$|\Phi_{\nu,\omega}\rangle = \varsigma_{\nu,\omega}\,|\phi_{\nu,\omega}\rangle + \sqrt{1-|\varsigma_{\nu,\omega}|^2}\,|\tilde{\phi}^\perp_{\nu,\omega}\rangle\,, \qquad \text{(B7)}$$

where $\varsigma_{\nu,\omega} = \langle\phi_{\nu,\omega}|\Phi_{\nu,\omega}\rangle$ and $|\tilde{\phi}^\perp_{\nu,\omega}\rangle$ is some state orthogonal to $|\phi_{\nu,\omega}\rangle$. Then, $\delta_{\nu,\omega}$ can be written as

$$\delta_{\nu,\omega} = |\sqrt{1-\epsilon_{\nu,\omega}}\varsigma_{\nu,\omega} + \sqrt{\epsilon_{\nu,\omega}}\sqrt{1-|\varsigma_{\nu,\omega}|^2}\langle\phi^\perp_{\nu,\omega}|\tilde{\phi}^\perp_{\nu,\omega}\rangle|. \qquad \text{(B8)}$$

In our particular case, $\varsigma_{\nu,\omega}$ depends on the parameter ξ defined in the main text. Specifically, $\varsigma_{\nu,\omega} = \xi$ when $\nu = \omega = \text{vac}$, $\varsigma_{\nu,\omega} = \sqrt{\xi}$ when either $\omega \neq \nu = \text{vac}$ or $\nu \neq \omega = \text{vac}$, and $\varsigma_{\nu,\omega} = 1$ otherwise. Thus, a lower bound on $\delta_{\nu,\omega}$ is straightforwardly given by

$$\delta^L_{\nu,\omega} = \sqrt{1-\epsilon_{\nu,\omega}}\varsigma_{\nu,\omega} - \sqrt{\epsilon_{\nu,\omega}}\sqrt{1-|\varsigma_{\nu,\omega}|^2}. \qquad \text{(B9)}$$

APPENDIX C: SECURITY AGAINST COHERENT ATTACKS

Here we briefly show that the analysis presented in the main text can be used to guarantee security against coherent attacks. For this, note that, for a protocol with N rounds, Eq. (11) is still valid for each particular round

$n = 1, \ldots, N$. Also, let us define $p_{\mathcal{K}}$ to be the probability that a round is selected for key generation. That is, this is the probability that in a successful round Alice and Bob *neither select the vacuum states nor the round is chosen to estimate the bit error rate or the phase error rate*. The probability $p_{\mathcal{K}}$ can be included as a factor on the right-hand side of Eq. (11), so we obtain an upper bound on the probability that the round n is used for key generation and a phase error occurs, namely $\Gamma_{\mathcal{K}}^{(n)}$. That is,

$$\Gamma_{\mathcal{K}}^{(n)} = p_{\mathcal{K}} \Gamma_n^U = p_{\mathcal{K}} G_+(\Gamma_{\mathrm{ref},n}^U, \delta_{\mathrm{vir}}^L), \tag{C1}$$

where Γ_n^U ($\Gamma_{\mathrm{ref},n}^U$) is an upper bound on the phase error probability of the actual (reference) states in round n. Then, by using Jensen's inequality [66], we obtain

$$\begin{aligned}\frac{1}{N}\sum_n \Gamma_{\mathcal{K}}^{(n)} &= \frac{1}{N}\sum_n p_{\mathcal{K}} G_+(\Gamma_{\mathrm{ref},n}^U, \delta_{\mathrm{vir}}^L)\\ &\le p_{\mathcal{K}} G_+\left(\frac{1}{N}\sum_n \Gamma_{\mathrm{ref},n}^U, \delta_{\mathrm{vir}}^L\right),\end{aligned} \tag{C2}$$

due to the concavity of G_+ with respect to its first element. Now, we can take advantage of the fact that the function Γ_{ref}^U given in Eq. (B6) is also concave with respect to $Y_{\nu,\omega}$, which, for a particular round n, we denote as $Y_{\nu,\omega}^n$, and again apply Jensen's inequality, now to $(1/N)\sum_n \Gamma_{\mathrm{ref},n}^U$, so we have

$$\begin{aligned}\frac{1}{N}\sum_n \Gamma_{\mathrm{ref},n}^U &= \frac{1}{N}\sum_n \sum_{\substack{\nu,\omega\\ f_{\nu,\omega}>0}} f_{\nu,\omega} G_+\left(\frac{\tilde{Y}^n_{\nu,\omega,\mathcal{T}}}{p_{\nu,\omega} p_{\mathcal{T}|\nu,\omega}}, \delta^L_{\nu,\omega}\right)\\ &\quad + \frac{1}{N}\sum_n \sum_{\substack{\nu,\omega\\ f_{\nu,\omega}<0}} f_{\nu,\omega} G_-\left(\frac{\tilde{Y}^n_{\nu,\omega,\mathcal{T}}}{p_{\nu,\omega} p_{\mathcal{T}|\nu,\omega}}, \delta^L_{\nu,\omega}\right)\\ &\le \sum_{\substack{\nu,\omega\\ f_{\nu,\omega}>0}} f_{\nu,\omega} G_+\left(\frac{\sum_n \tilde{Y}^n_{\nu,\omega,\mathcal{T}}}{N p_{\nu,\omega} p_{\mathcal{T}|\nu,\omega}}, \delta^L_{\nu,\omega}\right)\\ &\quad + \sum_{\substack{\nu,\omega\\ f_{\nu,\omega}<0}} f_{\nu,\omega} G_-\left(\frac{\sum_n \tilde{Y}^n_{\nu,\omega,\mathcal{T}}}{N p_{\nu,\omega} p_{\mathcal{T}|\nu,\omega}}, \delta^L_{\nu,\omega}\right),\end{aligned} \tag{C3}$$

where $\tilde{Y}^n_{\nu,\omega,\mathcal{T}} := Y^n_{\nu,\omega} p_{\nu,\omega} p_{\mathcal{T}|\nu,\omega}$ is the joint probability that Alice and Bob send $|\Psi_{\nu,\omega}\rangle$, Charles announces a successful event in round n, and the round is used for parameter estimation, $p_{\nu,\omega} = p_\nu p_\omega$, and $p_{\mathcal{T}|\nu,\omega}$ is the conditional probability that the round is used for parameter estimation given that Alice and Bob send $|\nu\rangle$ and $|\omega\rangle$, respectively. Note that $p_{\mathcal{T}|\nu,\omega} = 1$ if any of Alice's or Bob's states is the vacuum state. Also, we have $p_{\mathcal{K}} = 1 - \sum_{\nu,\omega} p_{\nu,\omega} p_{\mathcal{T}|\nu,\omega}$.

By combining Eqs. (C2) and (C3), one arrives at the bound

$$\begin{aligned}\sum_n \Gamma_{\mathcal{K}}^{(n)} \le N p_{\mathcal{K}} G_+\Bigg[&\sum_{\substack{\nu,\omega\\ f_{\nu,\omega}>0}} f_{\nu,\omega} G_+\left(\frac{\sum_n \tilde{Y}^n_{\nu,\omega,\mathcal{T}}}{N p_{\nu,\omega} p_{\mathcal{T}|\nu,\omega}}, \delta^L_{\nu,\omega}\right)\\ &+ \sum_{\substack{\nu,\omega\\ f_{\nu,\omega}<0}} f_{\nu,\omega} G_-\left(\frac{\sum_n \tilde{Y}^n_{\nu,\omega,\mathcal{T}}}{N p_{\nu,\omega} p_{\mathcal{T}|\nu,\omega}}, \delta^L_{\nu,\omega}\right), \delta^L_{\mathrm{vir}}\Bigg].\end{aligned} \tag{C4}$$

Importantly, the probability $\tilde{Y}^n_{\nu,\omega,\mathcal{T}}$ could depend on all the available information up to the nth round. This means that, with a negligible probability of failure for $N \to \infty$, one can estimate the sums $\sum_n \tilde{Y}^n_{\nu,\omega,\mathcal{T}}$ from the observed number of successful events within the parameter estimation rounds where Alice and Bob send $|\Psi_{\nu,\omega}\rangle$, namely $\tilde{N}_{\nu,\omega,\mathcal{T}}$, by using Azuma's inequality [67] or Kato's inequality [68]. That is, we have $\tilde{N}_{\nu,\omega,\mathcal{T}} \approx \sum_n \tilde{Y}^n_{\nu,\omega,\mathcal{T}}$. Moreover, Kato's inequality has already been demonstrated to offer very tight estimations for reasonable values of N [69], so we expect that the performance of the protocol should not be severely affected by finite-key effects.

Finally, it is possible to obtain an estimation on the number of phase errors, $\tilde{N}_{\mathrm{ph}}$, from the sum $\sum_n \Gamma_{\mathcal{K}}^{(n)}$ by again applying Azuma's or Kato's inequality. That is, we have $\tilde{N}_{\mathrm{ph}} \approx \sum_n \Gamma_{\mathcal{K}}^{(n)}$ with negligible probability of failure when $N \to \infty$.

APPENDIX D: OPTIMAL AMPLITUDE α

Here we show in Fig. 3, for completeness, the optimized values of the parameter α corresponding to the simulations shown in Fig. 2 of the main text.

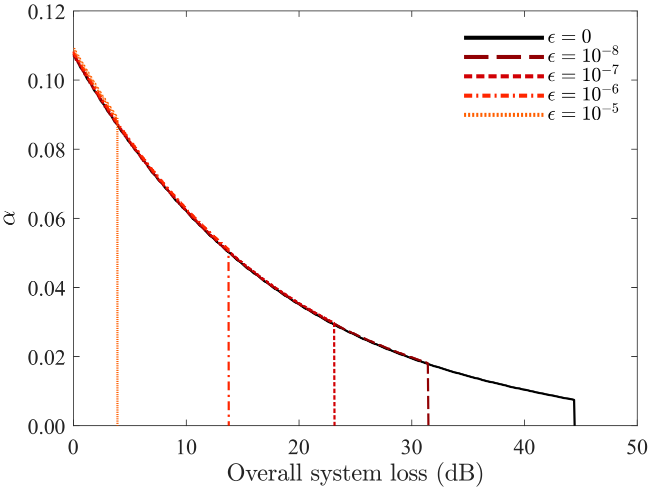

FIG. 3. Optimal value of α corresponding to the simulations shown in Fig. 2 in the main text.

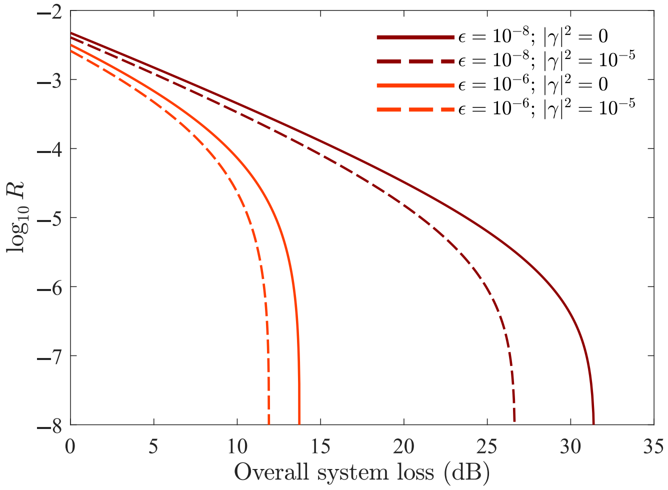

FIG. 4. Comparison between the ideal scenario where the third reference state $|\gamma\rangle$ used by Alice and Bob is a perfect vacuum state ($|\gamma|^2 = 0$) and the case where, instead, such a state is a weak coherent state ($|\gamma|^2 = 10^{-5}$). As can be observed, the performance of the protocol is similar in both cases.

APPENDIX E: NONVACUUM INTENSITY

Here we illustrate the effect that the use of imperfect vacuum states has on the performance of the protocol. For this, we consider the secret key rate that Alice and Bob would obtain when they use the set of states $\{|\alpha\rangle, |-\alpha\rangle, |\gamma\rangle\}$ with $\gamma \in \mathbb{R}$. That is, this set of states is used in the simulations to calculate the experimental probabilities $Y_{\nu,\omega}$ as well as to define the set of reference states $\{|\alpha\rangle, |-\alpha\rangle, |\gamma'\rangle\}$, with $|\gamma'\rangle$ the projection of $|\gamma\rangle$ onto the qubit space spanned by $\{|\alpha\rangle, |-\alpha\rangle\}$. Note, however, that any $|\gamma|^2 > 0$ could also be treated as an imperfection and thus it could be incorporated into the security proof by properly choosing the parameters $\epsilon_{\nu,\omega}$. The parameter α is optimized for each value of the overall system loss, and for illustration purposes, we evaluate two cases for the intensity $|\gamma|^2$: 0 and 10^{-5}. As one can see from Fig. 4, the performance is very similar in both cases, only slightly lower when $|\gamma|^2 = 10^{-5}$.

[1] H.-K. Lo, M. Curty, and K. Tamaki, Secure quantum key distribution, Nat. Photonics **8**, 595 (2014).

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

[3] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, and C. Ottaviani *et al.*, Advances in quantum cryptography, Adv. Opt. Photon. **12**, 1012 (2020).

[4] H.-L. Yin, T.-Y. Chen, Z.-W. Yu, H. Liu, L.-X. You, Y.-H. Zhou, S.-J. Chen, Y. Mao, M.-Q. Huang, and W.-J. Zhang *et al.*, Measurement Device Independent Quantum key Distribution Over 404 km Optical Fibre, Phys. Rev. Lett. **117**, 190501 (2016).

[5] A. Boaron, G. Boso, D. Rusca, C. Vulliez, C. Autebert, M. Caloz, M. Perrenoud, G. Gras, F. Bussières, and M.-J. Li *et al.*, Secure Quantum key Distribution Over 421 km of Optical Fiber, Phys. Rev. Lett. **121**, 190502 (2018).

[6] J.-P. Chen, C. Zhang, Y. Liu, C. Jiang, W. Zhang, X.-L. Hu, J.-Y. Guan, Z.-W. Yu, H. Xu, and J. Lin *et al.*, Sending-Or-Not-Sending with Independent Lasers: Secure Twin-Field Quantum key Distribution Over 509 km, Phys. Rev. Lett. **124**, 070501 (2020).

[7] S.-K. Liao, W.-Q. Cai, W.-Y. Liu, L. Zhang, Y. Li, J.-G. Ren, J. Yin, Q. Shen, Y. Cao, and Z.-P. Li *et al.*, Satellite-to-ground quantum key distribution, Nature **549**, 43 (2017).

[8] H. Takenaka, A. Carrasco-Casado, M. Fujiwara, M. Kitamura, M. Sasaki, and M. Toyoshima, Satellite-to-ground quantum-limited communication using a 50-kg-class microsatellite, Nat. Photonics **11**, 502 (2017).

[9] J. Yin, Y.-H. Li, S.-K. Liao, M. Yang, Y. Cao, L. Zhang, J.-G. Ren, W.-Q. Cai, W.-Y. Liu, and S.-L. Li *et al.*, Entanglement-based secure quantum cryptography over 1,120 kilometres, Nature **582**, 501 (2020).

[10] M. Peev, C. Pacher, R. Alléaume, C. Barreiro, J. Bouda, W. Boxleitner, T. Debuisschert, E. Diamanti, M. Dianati, and J. Dynes *et al.*, The SECOQC quantum key distribution network in Vienna, New J. Phys. **11**, 075001 (2009).

[11] M. Sasaki, M. Fujiwara, H. Ishizuka, W. Klaus, K. Wakui, M. Takeoka, S. Miki, T. Yamashita, Z. Wang, and A. Tanaka *et al.*, Field test of quantum key distribution in the Tokyo QKD network, Opt. Express **19**, 10387 (2011).

[12] J. Dynes, A. Wonfor, W.-S. Tam, A. Sharpe, R. Takahashi, M. Lucamarini, A. Plews, Z. Yuan, A. Dixon, and J. Cho *et al.*, Cambridge quantum network, Npj Quantum Inf. **5**, 1 (2019).

[13] D. Mayers and A. Yao, in *Proc. 39th Annual Symposium on Foundations of Computer Science* (IEEE, Los Alamitos, 1998), p. 503.

[14] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[15] U. Vazirani and T. Vidick, Fully Device-Independent Quantum key Distribution, Phys. Rev. Lett. **113**, 140501 (2014).

[16] J. S. Bell, On the einstein podolsky rosen paradox, Physics **1**, 195 (1964).

[17] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[18] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. Vermeulen, R. N. Schouten, and C. Abellán *et al.*, Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres, Nature **526**, 682 (2015).

[19] V. Zapatero and M. Curty, Long-distance device-independent quantum key distribution, Sci. Rep. **9**, 1 (2019).

[20] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, Trojan-horse attacks on quantum-key-distribution systems, Phys. Rev. A **73**, 022320 (2006).

[21] N. Jain, E. Anisimova, I. Khan, V. Makarov, C. Marquardt, and G. Leuchs, Trojan-horse attacks threaten the security of

practical quantum cryptography, New J. Phys. **16**, 123030 (2014).

[22] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. Yuan, and A. J. Shields, Practical Security Bounds against the Trojan-Horse Attack in Quantum key Distribution, Phys. Rev. X **5**, 031030 (2015).

[23] C. Kurtsiefer, P. Zarda, S. Mayer, and H. Weinfurter, The breakdown flash of silicon avalanche photodiodes-back door for eavesdropper attacks?, J. Mod. Opt. **48**, 2039 (2001).

[24] A. Meda, I. P. Degiovanni, A. Tosi, Z. Yuan, G. Brida, and M. Genovese, Quantifying backflash radiation to prevent zero-error attacks in quantum key distribution, Light: Sci. Appl. **6**, e16261 (2017).

[25] P. V. P. Pinheiro, P. Chaiwongkhot, S. Sajeed, R. T. Horn, J.-P. Bourgoin, T. Jennewein, N. Lütkenhaus, and V. Makarov, Eavesdropping and countermeasures for backflash side channel in quantum cryptography, Opt. Express **26**, 21020 (2018).

[26] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Hacking commercial quantum cryptography systems by tailored bright illumination, Nat. Photonics **4**, 686 (2010).

[27] N. Jain, B. Stiller, I. Khan, D. Elser, C. Marquardt, and G. Leuchs, Attacks on practical quantum key distribution systems (and how to prevent them), Contemp. Phys. **57**, 3 (2016).

[28] H.-K. Lo, M. Curty, and B. Qi, Measurement-Device-Independent Quantum key Distribution, Phys. Rev. Lett. **108**, 130503 (2012).

[29] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, Overcoming the rate–distance limit of quantum key distribution without quantum repeaters, Nature **557**, 400 (2018).

[30] X.-B. Wang, Z.-W. Yu, and X.-L. Hu, Twin-field quantum key distribution with large misalignment error, Phys. Rev. A **98**, 062323 (2018).

[31] M. Curty, K. Azuma, and H.-K. Lo, Simple security proof of twin-field type quantum key distribution protocol, Njp Quantum Inf. **5**, 64 (2019).

[32] X. Ma, P. Zeng, and H. Zhou, Phase-Matching Quantum key Distribution, Phys. Rev. X **8**, 031043 (2018).

[33] C. Cui, Z.-Q. Yin, R. Wang, W. Chen, S. Wang, G.-C. Guo, and Z.-F. Han, Twin-Field Quantum key Distribution Without Phase Postselection, Phys. Rev. Appl. **11**, 034053 (2019).

[34] K. Tamaki, H.-K. Lo, W. Wang, and M. Lucamarini, Information theoretic security of quantum key distribution overcoming the repeaterless secret key capacity bound, preprint arXiv:1805.05511 (2018).

[35] J. Lin and N. Lütkenhaus, Simple security analysis of phase-matching measurement-device-independent quantum key distribution, Phys. Rev. A **98**, 042332 (2018).

[36] M. Pereira, M. Curty, and K. Tamaki, Quantum key distribution with flawed and leaky sources, Npj Quantum Inf. **5**, 1 (2019).

[37] M. Pereira, G. Kato, A. Mizutani, M. Curty, and K. Tamaki, Quantum key distribution with correlated sources, Sci. Adv. **6**, eaaz4487 (2020).

[38] K. Tamaki, M. Curty, G. Kato, H.-K. Lo, and K. Azuma, Loss-tolerant quantum cryptography with imperfect sources, Phys. Rev. A **90**, 052314 (2014).

[39] Z. Tang, K. Wei, O. Bedroya, L. Qian, and H.-K. Lo, Experimental measurement-device-independent quantum key distribution with imperfect sources, Phys. Rev. A **93**, 042308 (2016).

[40] Z. Cao, Z. Zhang, H.-K. Lo, and X. Ma, Discrete-phase-randomized coherent state source and its application in quantum key distribution, New J. Phys. **77**, 053014 (2015).

[41] K. Tamaki, M. Curty, and M. Lucamarini, Decoy-state quantum key distribution with a leaky source, New J. Phys. **18**, 065008 (2016).

[42] W. Wang, K. Tamaki, and M. Curty, Finite-key security analysis for quantum key distribution with leaky sources, New J. Phys. **20**, 083027 (2018).

[43] K. Yoshino, M. Fujiwara, K. Nakata, T. Sumiya, T. Sasaki, M. Takeoka, M. Sasaki, A. Tajima, M. Koashi, and A. Tomita, Quantum key distribution with an efficient countermeasure against correlated intensity fluctuations in optical pulses, Npj Quantum Inf. **4**, 8 (2018).

[44] F. Xu, M. Curty, B. Qi, and H.-K. Lo, Practical aspects of measurement-device-independent quantum key distribution, New J. Phys. **15**, 113007 (2013).

[45] Y. Liu, Z.-W. Yu, W. Zhang, J.-Y. Guan, J.-P. Chen, C. Zhang, X.-L. Hu, H. Li, C. Jiang, and J. Lin *et al.*, Experimental Twin-Field Quantum key Distribution through Sending or not Sending, Phys. Rev. Lett. **123**, 100505 (2019).

[46] F. Grasselli, A. Navarrete, and M. Curty, Asymmetric twin-field quantum key distribution, New J. Phys. **21**, 113032 (2019).

[47] W. Wang, F. Xu, and H.-K. Lo, Asymmetric Protocols for Scalable High-Rate Measurement-Device-Independent Quantum key Distribution Networks, Phys. Rev. X **9**, 041012 (2019).

[48] W. Wang and H.-K. Lo, Simple method for asymmetric twin-field quantum key distribution, New J. Phys. **22**, 013020 (2020).

[49] X. Zhong, W. Wang, L. Qian, and H.-K. Lo, Proof-of-principle experimental demonstration of twin-field quantum key distribution over optical channels with asymmetric losses, Npj Quantum Inf. **7**, 8 (2021).

[50] A. Vakhitov, V. Makarov, and D. R. Hjelme, Large pulse attack as a method of conventional optical eavesdropping in quantum cryptography, J. Mod. Opt. **48**, 2023 (2001).

[51] M. Koashi, Simple security proof of quantum key distribution based on complementarity, New J. Phys. **11**, 045018 (2009).

[52] M. Minder, M. Pittaluga, G. Roberts, M. Lucamarini, J. Dynes, Z. Yuan, and A. Shields, Experimental quantum key distribution beyond the repeaterless secret key capacity, Nat. Photonics **13**, 334 (2019).

[53] S. Wang, D.-Y. He, Z.-Q. Yin, F.-Y. Lu, C.-H. Cui, W. Chen, Z. Zhou, G.-C. Guo, and Z.-F. Han, Beating the Fundamental Rate-Distance Limit in a Proof-Of-Principle Quantum key Distribution System, Phys. Rev. X **9**, 021046 (2019).

[54] X. Zhong, J. Hu, M. Curty, L. Qian, and H.-K. Lo, Proof-Of-Principle Experimental Demonstration of Twin-Field Type Quantum key Distribution, Phys. Rev. Lett. **123**, 100506 (2019).

[55] M. Pittaluga, M. Minder, M. Lucamarini, M. Sanzaro, R. I. Woodward, M.-J. Li, Z. Yuan, and A. J. Shields, 600 km

repeater-like quantum communications with dual-band stabilisation, preprint arXiv:2012.15099 (2020).

[56] M. Curty and T. Moroder, Heralded-qubit amplifiers for practical device-independent quantum key distribution, Phys. Rev. A **84**, 010304 (2011).

[57] W.-Y. Hwang, Quantum key Distribution with High Loss: Toward Global Secure Communication, Phys. Rev. Lett. **91**, 057901 (2003).

[58] H.-K. Lo, X. Ma, and K. Chen, Decoy State Quantum key Distribution, Phys. Rev. Lett. **94**, 230504 (2005).

[59] X.-B. Wang, Beating the Photon-Number-Splitting Attack in Practical Quantum Cryptography, Phys. Rev. Lett. **94**, 230503 (2005).

[60] A. Huang, A. Navarrete, S.-H. Sun, P. Chaiwongkhot, M. Curty, and V. Makarov, Laser-Seeding Attack in Quantum key Distribution, Phys. Rev. Appl. **12**, 064043 (2019).

[61] S.-H. Sun, M. Gao, M.-S. Jiang, C.-Y. Li, and L.-M. Liang, Partially random phase attack to the practical two-way quantum-key-distribution system, Phys. Rev. A **85**, 032304 (2012).

[62] S.-H. Sun, F. Xu, M.-S. Jiang, X.-C. Ma, H.-K. Lo, and L.-M. Liang, Effect of source tampering in the security of quantum cryptography, Phys. Rev. A **92**, 022304 (2015).

[63] K. Tamaki, H.-K. Lo, C.-H. F. Fung, and B. Qi, Phase encoding schemes for measurement-device-independent quantum key distribution with basis-dependent flaw, Phys. Rev. A **85**, 042307 (2012).

[64] F. Xu, H. Xu, and H.-K. Lo, Protocol choice and parameter optimization in decoy-state measurement-device-independent quantum key distribution, Phys. Rev. A **89**, 052333 (2014).

[65] W. Wang, K. Tamaki, and M. Curty, Measurement-device-independent quantum key distribution with leaky sources, Sci. Rep. **11**, 1678 (2021).

[66] J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, Acta Mathematica **30**, 175 (1906).

[67] K. Azuma, Weighted sums of certain dependent random variables, Tohoku Math. J. **19**, 357 (1967).

[68] G. Kato, Concentration inequality using unconfirmed knowledge, preprint arXiv:2002.04357 (2020).

[69] G. Currás-Lorenzo, Á. Navarrete, K. Azuma, G. Kato, M. Curty, and M. Razavi, Tight finite-key security for twin-field quantum key distribution, Npj Quantum Inf. **7**, 1 (2021).