npj | Quantum Information

www.nature.com/npjqi

ARTICLE OPEN

Quantum key distribution with flawed and leaky sources

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

In theory, quantum key distribution (QKD) allows secure communications between two parties based on physical laws. However, most of the security proofs of QKD today make unrealistic assumptions and neglect many relevant device imperfections. As a result, they cannot guarantee the security of the practical implementations. Recently, the loss-tolerant protocol (K. Tamaki et al., Phys. Rev. A, 90, 052314, 2014) was proposed to make QKD robust against state preparation flaws. This protocol relies on the emission of qubit systems, which, unfortunately, is difficult to achieve in practice. In this work, we remove such qubit assumption and generalise the loss-tolerant protocol to accommodate multiple optical modes in the emitted signals. These multiple optical modes could arise, e.g., from Trojan horse attacks and/or device imperfections. Our security proof determines some dominant device parameter regimes needed for achieving secure communication and, therefore, it can serve as a guideline to characterise QKD transmitters. Furthermore, we compare our approach with that of H.-K. Lo et al. (Quantum Inf. Comput., 7, 431–458, 2007) and identify which method provides the highest secret key generation rate as a function of the device imperfections. Our work constitutes an important step towards the best practical and secure implementation for QKD.

npj Quantum Information (2019)5:62 ; https://doi.org/10.1038/s41534-019-0180-9

INTRODUCTION

Quantum key distribution (QKD)[1–3] enables two distant parties, Alice and Bob, to share a common secret key that can be used to encrypt and decrypt messages. In theory, QKD can offer information-theoretic security based on the laws of physics. In practice, however, it does not, because typical security proofs of QKD require assumptions that are not actually met by the practical implementations, as they usually ignore many experimental device imperfections. This discrepancy between the theory and the practice of QKD has been evidenced by many quantum hacking attacks, especially by those that exploit flaws in the detectors of QKD systems.[4,5] Fortunately, the proposal of measurement-device-independent QKD (MDI-QKD)[6] can solve all security loopholes in the measurement unit and, therefore, Eve cannot take advantage of detector side channels to learn information about the key. Furthermore, MDI-QKD can be implemented experimentally using standard optical components.[7–12] Therefore, to guarantee implementation security we now need to focus on how to secure the source in QKD.

Ideally, the sending devices are single-photon sources and the encoding of the light pulses is executed perfectly, without any state preparation flaw (SPF). However, none of these two conditions are met experimentally, as all devices have inherent deficiencies. The decoy-state method[13–15] was proposed to replace single-photon sources with coherent light sources. Also, by using the Gottesman–Lo–Lütkenhaus–Preskill (GLLP) security analysis[16] the problem with SPFs is fixed. The main drawback of this last approach is that the resulting secret key rate is poor and fragile against channel loss. This is because it assumes the worst-case scenario in which Eve could enhance the signals' flaws through channel loss, which significantly decreases the performance of the QKD scheme.

Recently, a protocol that is loss-tolerant (LT) to SPFs has been proposed[17] to address the limitation of the GLLP analysis. The LT protocol employs only three states and takes into account modulation errors due to an imperfect phase modulator (PM). Remarkably, by using a different phase error estimation technique involving the use of the basis mismatched events, the secret key rate of the LT protocol remains almost unchanged even if the SPFs increase. In fact, the maximum transmission distance for a QKD system in fiber so far has been recently achieved using this protocol,[18] which shows that the LT protocol is highly practical. Its main weakness, however, is the assumption that the single-photon signals sent by Alice are qubits, which is difficult to guarantee in practice. For instance, if Eve conducts a Trojan horse attack (THA)[19–23] against the source, this assumption can be violated. In a THA, Eve sends bright light into Alice's PM and obtains information about the encoding by measuring the back-reflected light that exits Alice's lab. Moreover, an optical mode of the light pulse emitted by Alice could be dependent on the value of the phase modulation, which means that a sent single-photon pulse might not be a qubit (we call this imperfection the non-qubit assumption). That is, Alice's setting choice information could be encoded in other degrees of freedom of the emitted light, and this spontaneous leakage of information results in a higher-dimensional sending state.

This work aims to reduce the big gap between the theory and the practice of QKD by generalising the LT protocol such that it can include typical imperfections in the sending device. To be precise, and in contrast to,[17] here we remove the qubit assumption and include the effect of side-channels by considering the mode dependency of the PM and THAs. Moreover, like in ref. [17], we also include in the analysis SPFs in a single-mode qubit subspace. Therefore, our analysis covers dominant imperfections that a source device has, allowing the use of a much wider class of

[1]Escuela de Ingeniería de Telecomunicación, Department of Signal Theory and Communications, University of Vigo, Vigo E-36310, Spain and [2]Graduate School of Science and Engineering for Research, University of Toyama, Gofuku 3190, Toyama 930-8555, Japan
Correspondence: Margarida Pereira (mpereira@com.uvigo.es)

Received: 13 February 2019 Accepted: 1 July 2019
Published online: 26 July 2019

Published in partnership with The University of New South Wales

npj nature partner journals

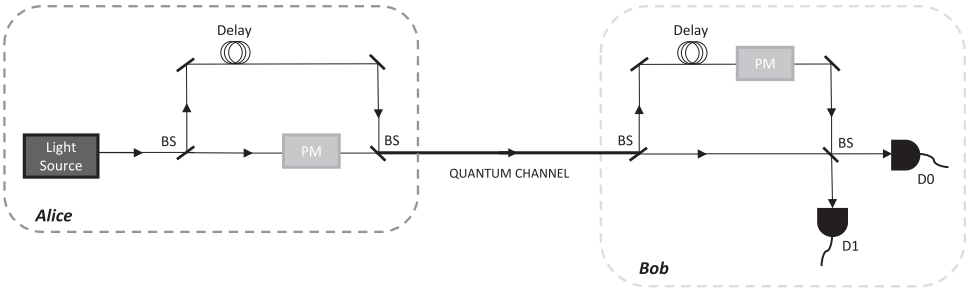

Fig. 1 Each single-photon pulse emitted by Alice's source goes through a 50:50 beamsplitter (BS) and is decomposed into the reference and the signal pulses. The reference pulse travels through the longer arm of Alice's Mach–Zehnder interferometer. To perform the encoding, she uses a PM that applies a phase shift to the signal pulse. The two pulses are recombined at the second 50:50 BS, sent through the quantum channel and then received in Bob's lab. On reception, they are split by a 50:50 BS and Bob applies a phase shift on the reference and signal pulses in the upper arm of his Mach–Zehnder interferometer. These pulses then interfere with the pulses that travelled through the shorter arm of the interferometer at the second 50:50 BS. Bob can then detect click events corresponding to photons choosing the shortest arm in Alice's interferometer and the longest one in Bob's, and to the opposite, by using two detectors, D0 and D1, which correspond to obtaining bit value 0 and 1, respectively

imperfect devices in a secure manner. Our generalised LT protocol can be applied to any multi-mode scenario as long as the states of the emitted signals are independently and identically distributed (I.I.D.), namely, this proof does not consider correlations between the sending signals. However, we remark that recent results reported in ref. [24] imply that our analysis could also accommodate correlations between the signals which are independent of Alice's setting choice. Furthermore, we emphasise that the basic idea is rather general and can be applied to many other QKD protocols such as, for instance, the six-state protocol,[25] distributed-phase-reference protocols[26–28] and MDI-QKD.[6] In simple terms, it is a formalism to estimate the phase error rate of a QKD protocol by evaluating the transmission rates of some virtual states with the help of the state structure of Alice's signals (see Eq. (1) below). We also emphasise that our method does not require a complete characterisation of the side channels, which significantly simplifies the experiments for characterising the source. Using this formalism, we can quantify the device parameters required to ensure secure communications with flawed and leaky sources.

In addition, we investigate how Lo–Preskill's (LP) security analysis[29] behaves in the presence of the same device's imperfections and, by using imperfectly characterised states, we compare it with our generalised LT protocol. As a result, we determine which security proof provides a higher secret key rate as a function of the device parameters. These parameters are essential for experimentalists to produce and to calibrate the transmitting devices, and therefore our work can be used as a guideline for securing the source in the presence of multi-mode signals.

RESULTS

Description of the protocol

We shall assume, for simplicity, that Alice's lab has a single-photon source. However, we emphasise that our analysis can also be applied to the case where Alice emits phase-randomised weak coherent pulses. In this latter case, Alice can use the decoy-state method[13–15] to estimate all the quantities corresponding to the single-photon pulses that are needed to apply our method. Below we focus on the case where Alice has at her disposal single-photon sources only because the study with phase-randomised weak coherent pulses, together with decoy states, results in an unnecessarily cumbersome analysis. Figure 1 shows the QKD setup (see Section I of the Supplementary Material for a detailed description of the actual protocol). Next, we describe the assumptions we make on Alice's and Bob's devices.

Assumptions on Alice's device

In this work, we consider the asymptotic scenario where Alice sends Bob an infinite number of pulses. Our formalism is valid for any source that emits pulses whose quantum state is of the form

$$\left|\Phi_{j\beta}\right\rangle_{\mathrm{BE}} = a_{j\beta}\left|\phi_{j\beta}\right\rangle_{\mathrm{BE}} + b_{j\beta}\left|\phi_{j\beta}^{\perp}\right\rangle_{\mathrm{BE}}, \tag{1}$$

with $|a_{j\beta}|^2 + |b_{j\beta}|^2 = 1$, where $j \in \{0, 1\}$ and $\beta \in \{Z, X\}$ are Alice's bit value and basis choices, respectively. As in the LT analysis introduced in ref. [17], we consider a three-state protocol where Alice selects $j\beta \in \{0Z, 1Z, 0X\}$. Furthermore, in Eq. (1), we assume that $|\phi_{j\beta}\rangle_{\mathrm{BE}}$ is a pure state in a single-mode qubit space, where BE stands for Bob's and Eve's systems due to a potential THA. For instance, $|\phi_{j\beta}\rangle_{\mathrm{BE}}$ could be of the form $|\phi_{j\beta}\rangle_{\mathrm{BE}} = |\omega_{j\beta}\rangle_{\mathrm{B}} \otimes |\epsilon\rangle_{\mathrm{E}}$, where Eve's system does not depend on $j\beta$ and $|\omega_{j\beta}\rangle_{\mathrm{B}}$ is a qubit state. The state $\left|\phi_{j\beta}^{\perp}\right\rangle_{\mathrm{BE}}$, on the other hand, corresponds to any state outside of the single-mode qubit space, including the state of a side channel, and it is in an Hilbert space orthogonal to $|\varphi_{j\beta}\rangle_{\mathrm{BE}}$. We note that the form of the pure state given by Eq. (1) is the most general I.I.D. state. Indeed, this equation simply decomposes a state in a given Hilbert space into a direct sum of two states in different Hilbert spaces, which can always be done. One of these states is in a qubit space and the other one is in any complementary Hilbert space. That is, any pure state can be written in the form given by Eq. (1). In addition, we further assume that, similar to that in ref. [17], the states $|\phi_{j\beta}\rangle_{\mathrm{BE}}$ in Eq. (1) form a triangle in the Bloch sphere and we set their Y-components to be the same by choosing the y-axis appropriately. This assumption is required to ensure that Alice is sending essentially three different states rather than one or two states. Importantly, we note that by introducing an ancilla system for Alice to purify the state, our formalism is also valid for a mixed state in a single-mode qubit space, as shown in section 'Security proof against coherent attacks'.

The state structure in Eq. (1) means that the inner product $\left\langle \phi_{j\beta}|\phi_{j'\beta'}^{\perp}\right\rangle_{\mathrm{BE}}$ for all j, j', β and β' is always zero. Also, depending on Alice's knowledge about the state given in Eq. (1), she might have to consider the worst-case scenario, i.e., $\left\langle \phi_{j\beta}^{\perp}|\phi_{j'\beta'}^{\perp}\right\rangle_{\mathrm{BE}} = 0$ for any combination of (j, β) and (j', β'). This means that, complete information about $\left|\phi_{j\beta}^{\perp}\right\rangle$ is not required, which significantly simplifies the experiments for characterising the source. On the other hand, if Alice knows some structure of the side channel she should fully exploit it and lower bound $\left\langle \phi_{j\beta}^{\perp}|\phi_{j'\beta'}^{\perp}\right\rangle_{\mathrm{BE}}$. For example, if she knows that the side channel is associated with the polarisation state of the single-mode qubit, then the worst-case

scenario does not apply, i.e., $\left\langle \phi_{j\beta}^{\perp} | \phi_{j'\beta'}^{\perp} \right\rangle_{\mathrm{BE}} \neq 0$, as it is impossible for three states to be orthogonal to each other given the two-dimensionality of polarisation. This way, our formalism can readily take into account the available information.

We remark that, to apply the procedure introduced below we only need to determine the coefficients $a_{j\beta}$ and $b_{j\beta}$, and the qubit state, but it is not necessary to completely characterise the quantum information of the side channel, $\left|\phi_{j\beta}^{\perp}\right\rangle_{\mathrm{BE}}$. That is, our characterisation seems to be rather simple and there is no need to perform further detailed characterisations. Nonetheless, the better Alice and Bob know the state given in Eq. (1), the better the resulting performance, as explained later in this section. An experimental procedure to perform this estimation is out of the scope of this paper; hence, we assume that these parameters are given.

Furthermore, our work can accommodate any SPF in the single-mode qubit space and one could also employ the techniques in refs. [24,30]. For example, we may select a case in which the states that Alice prepares can be expressed as

$$\frac{1}{\sqrt{2}}\left(|1\rangle_r|v\rangle_s + e^{\varphi_A+\delta\varphi_A/\pi}|v\rangle_r|1\rangle_s\right), \tag{2}$$

where $\delta(\geq 0)$ is the deviation of the phase modulation from the intended value φ_A and we define $|1\rangle_r|v\rangle_s = |0_Y\rangle$ and $|v\rangle_r|1\rangle_s = |1_Y\rangle$, where v stands for vacuum, $|1\rangle$ denotes a Fock state with one photon and the subscript r (s) corresponds to the reference (signal) pulse. In the case of the three-state protocol we have that $\varphi_A \in \{0,\ \pi,\ \pi/2\}$. Then, by using $|0_Z\rangle = (|0_Y\rangle + |1_Y\rangle)/\sqrt{2}$ and $|1_Z\rangle = (-|0_Y\rangle + |1_Y\rangle)/\sqrt{2}$, we obtain the following expressions for the three states in the single-mode qubit space:

$$\begin{aligned}
|\omega_{0Z}\rangle_{\mathrm{B}} &= |0_Z\rangle,\\
|\omega_{1Z}\rangle_{\mathrm{B}} &= -\sin\left(\tfrac{\delta}{2}\right)|0_Z\rangle + \cos\left(\tfrac{\delta}{2}\right)|1_Z\rangle,\\
|\omega_{0X}\rangle_{\mathrm{B}} &= \cos\left(\tfrac{\pi}{4}+\tfrac{\delta}{4}\right)|0_Z\rangle + \sin\left(\tfrac{\pi}{4}+\tfrac{\delta}{4}\right)|1_Z\rangle.
\end{aligned} \tag{3}$$

Therefore, our formalism can be used, for instance, when the information about Alice's choice of state is leaked and/or the optical mode depends on Alice's selection. This leakage from the source can occur spontaneously or with an active THA.

Assumptions on Bob's device

Bob receives the signal and reference pulses from Alice and measures them in a basis selected at random. More precisely, Bob's measurements are defined by the positive-operator valued measures (POVMs) $\{\hat{M}_{0\beta}, \hat{M}_{1\beta}, \hat{M}_f\}$, where $\hat{M}_{0\beta}$ ($\hat{M}_{1\beta}$) with $\beta \in \{X, Z\}$ corresponds to obtaining the bit value 0 (1) when Bob chooses the basis β and $\hat{M}_f$ corresponds to an inconclusive outcome. Importantly, $\hat{M}_f$ is assumed to be the same for the two bases. This means that the detection efficiencies are independent of Bob's measurement basis choice, which is required to prevent side-channel attacks exploiting channel loss.[4,5] It is noteworthy that this assumption is widely used in most security proofs and one of the simplest ways to circumvent such detector side-channel attacks is to use MDI-QKD, to which our technique also applies (see Section II in the Supplementary Material).

Security analysis

In order to prove the information-theoretic security of our protocol we use the complementary scenario introduced by Koashi.[31,32] For this, we first need to create an equivalent virtual protocol (see section 'Security proof against coherent attacks') concerning an observable conjugate to the key. The classical and quantum information available to Eve in the actual and virtual protocols are the same and therefore she cannot distinguish and behave differently between them. Hence, by proving the security in the virtual protocol we ensure the security of the actual protocol. In addition, from the virtual protocol we can determine the phase error rate, which quantifies the amount of information that is leaked to Eve and has to be removed in the privacy amplification step. In this section, we show how this last quantity is estimated by generalising the LT method.

As explained before, for simplicity we assume the asymptotic scenario where Alice sends Bob an infinite number of pulses. The asymptotic key rate for the single-photon signals can be expressed as

$$R \geq Y_Z[1 - h(e_X) - fh(e_Z)], \tag{4}$$

where Y_Z is the yield of the single photons in the Z basis, i.e., the joint probability of Alice emitting a single-photon in the Z basis and Bob detecting it with a measurement also in the Z basis. The function $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function and f is the error correction efficiency. The term e_X is the phase error rate and thus $h(e_X)$ is the cost of performing privacy amplification in order to remove the correlations between the corrected sifted key and Eve. The term e_Z is the bit error rate and $fh(e_Z)$ corresponds to the amount of syndrome information required to make Alice's and Bob's keys the same. The quantities Y_Z and e_Z in Eq. (4) can be directly obtained from an implementation of the experiment. Therefore, we are left with the estimation of the phase error rate. We do this below.

Estimation of the phase error rate

We assume that Alice prepares the states $|\Phi_{j\beta}\rangle_{\mathrm{BE}}$ as defined in Eq. (1). These states take into consideration the non-qubit assumption, a possible THA by Eve, and SPFs. In the virtual protocol (see section 'Security proof against coherent attacks' for further details), Alice prepares the following state in the Z basis:

$$\begin{aligned}
|\Psi_Z\rangle_{\mathrm{ABE}} = \tfrac{1}{\sqrt{2}}\Big[&|0_Z\rangle_A \otimes \left(a_{0Z}|\phi_{0Z}\rangle_{\mathrm{BE}} + b_{0Z}|\phi_{0Z}^{\perp}\rangle_{\mathrm{BE}}\right)\\
&+ |1_Z\rangle_A \otimes \left(a_{1Z}|\phi_{1Z}\rangle_{\mathrm{BE}} + b_{1Z}|\phi_{1Z}^{\perp}\rangle_{\mathrm{BE}}\right)\Big].
\end{aligned} \tag{5}$$

We then define the bit error rate as

$$e_Z = \frac{Y_{0Z,1Z}^{(Z)} + Y_{1Z,0Z}^{(Z)}}{Y_{0Z,0Z}^{(Z)} + Y_{1Z,0Z}^{(Z)} + Y_{0Z,1Z}^{(Z)} + Y_{1Z,1Z}^{(Z)}}, \tag{6}$$

where the yields $Y_{sZ,jZ}^{(Z)}$, with $s, j \in \{0, 1\}$, are the joint probabilities that Alice prepares the state $|\Psi_Z\rangle_{\mathrm{ABE}}$, Bob selects the Z basis and Alice (Bob) obtains the bit value j (s) when she (he) measures the system A (B) in the Z basis. It is noteworthy that the superscript (Z) in the yields represents the basis used in the state preparation, while the subscripts denote the bases employed in the measurements. These yields are directly observed in the experiment. Similarly, the phase error rate is defined as

$$e_X = \frac{Y_{0X,1X}^{(Z)vir} + Y_{1X,0X}^{(Z)vir}}{Y_{0X,0X}^{(Z)vir} + Y_{1X,0X}^{(Z)vir} + Y_{0X,1X}^{(Z)vir} + Y_{1X,1X}^{(Z)vir}}, \tag{7}$$

where $Y_{sX,jX}^{(Z)vir}$, with $s, j \in \{0, 1\}$, is the joint probability that Alice prepares the state $|\Psi_Z\rangle_{\mathrm{ABE}}$, she and Bob select the Z basis but both use the X basis for their measurements (rather than the selected Z basis) and Alice (Bob) obtains the bit value j (s). The phase error rate corresponds to the bit error in the virtual protocol. Also, we have that the denominator of e_X in Eq. (7) is equal to $Y_{0Z,0Z}^{(Z)} + Y_{1Z,0Z}^{(Z)} + Y_{0Z,1Z}^{(Z)} + Y_{1Z,1Z}^{(Z)}$, as by assumption the probability to obtain an inconclusive outcome associated to the operator $\hat{M}_f$ is the same for the both basis for any incoming state. This means that to estimate e_X we only need to calculate the virtual yields $Y_{0X,1X}^{(Z)vir}$ and $Y_{1X,0X}^{(Z)vir}$.

In the virtual protocol, after Alice measures the system A in Eq. (5) in the X basis, she sends Bob the (unnormalised) states:

$$\hat{\theta}_{\mathrm{BE},jX,vir} = \mathrm{Tr}_A\left[|j_X\rangle\langle j_X|_A \otimes \hat{1}_{\mathrm{BE}} |\Psi_Z\rangle\langle\Psi_Z|_{\mathrm{ABE}}\right], \tag{8}$$

where Tr_A is the partial trace over the virtual system A. Using Eqs. (5) and (8) we can calculate the unnormalised states sent by Alice for $j \in \{0, 1\}$ and obtain that $\hat{\theta}_{\mathrm{BE},jX,vir} = |\psi\rangle\langle\psi|_{\mathrm{BE},jX,vir}$ with

$$|\psi\rangle_{\mathrm{BE},jX,vir} = \tfrac{1}{2}\Big[a_{0Z}|\phi_{0Z}\rangle_{\mathrm{BE}} + b_{0Z}|\phi^{\perp}_{0Z}\rangle_{\mathrm{BE}} + (-1)^j\left(a_{1Z}|\phi_{1Z}\rangle_{\mathrm{BE}} + b_{1Z}|\phi^{\perp}_{1Z}\rangle_{\mathrm{BE}}\right)\Big]. \tag{9}$$

Writing Eq. (9) in terms of the states $|\gamma_{jX}\rangle_{\mathrm{BE}}$ and $\left|\gamma^{\perp}_{jX}\right\rangle_{\mathrm{BE}}$, defined below, we have that:

$$|\psi\rangle_{\mathrm{BE},jX,vir} = \tfrac{1}{2}\Big[\sqrt{|a_{0Z}|^2 + (-1)^j\left(a^*_{0Z}a_{1Z}\langle\phi_{0Z}|\phi_{1Z}\rangle_{\mathrm{BE}} + a_{0Z}a^*_{1Z}\langle\phi_{1Z}|\phi_{0Z}\rangle_{\mathrm{BE}}\right) + |a_{1Z}|^2}\,|\gamma_{jX}\rangle_{\mathrm{BE}} + \sqrt{|b_{0Z}|^2 + (-1)^j\left(b^*_{0Z}b_{1Z}\langle\phi^{\perp}_{0Z}|\phi^{\perp}_{1Z}\rangle_{\mathrm{BE}} + b_{0Z}b^*_{1Z}\langle\phi^{\perp}_{1Z}|\phi^{\perp}_{0Z}\rangle_{\mathrm{BE}}\right) + |b_{1Z}|^2}\left|\gamma^{\perp}_{jX}\right\rangle_{\mathrm{BE}}\Big], \tag{10}$$

where the normalised states $|\gamma_{jX}\rangle_{\mathrm{BE}}$ have the form

$$|\gamma_{jX}\rangle_{\mathrm{BE}} = \frac{a_{0Z}|\phi_{0Z}\rangle_{\mathrm{BE}} + (-1)^j a_{1Z}|\phi_{1Z}\rangle_{\mathrm{BE}}}{\sqrt{|a_{0Z}|^2 + (-1)^j\left(a^*_{0Z}a_{1Z}\langle\phi_{0Z}|\phi_{1Z}\rangle_{\mathrm{BE}} + a_{0Z}a^*_{1Z}\langle\phi_{1Z}|\phi_{0Z}\rangle_{\mathrm{BE}}\right) + |a_{1Z}|^2}}, \tag{11}$$

and the normalised states $\left|\gamma^{\perp}_{jX}\right\rangle_{\mathrm{BE}}$, which are orthogonal to $|\gamma_{jX}\rangle_{\mathrm{BE}}$, are given by

$$\left|\gamma^{\perp}_{jX}\right\rangle_{\mathrm{BE}} = \frac{b_{0Z}|\phi^{\perp}_{0Z}\rangle_{\mathrm{BE}} + (-1)^j b_{1Z}|\phi^{\perp}_{1Z}\rangle_{\mathrm{BE}}}{\sqrt{|b_{0Z}|^2 + (-1)^j\left(b^*_{0Z}b_{1Z}\langle\phi^{\perp}_{0Z}|\phi^{\perp}_{1Z}\rangle_{\mathrm{BE}} + b_{0Z}b^*_{1Z}\langle\phi^{\perp}_{1Z}|\phi^{\perp}_{0Z}\rangle_{\mathrm{BE}}\right) + |b_{1Z}|^2}}. \tag{12}$$

Note that, in Eq. (10), we have decomposed $|\psi\rangle_{\mathrm{BE},jX,vir}$ into a single-mode qubit $|\gamma_{jX}\rangle_{BE}$ and a state in any mode orthogonal to it, $\left|\gamma^{\perp}_{jX}\right\rangle_{\mathrm{BE}}$. This decomposition follows the definition provided in Eq. (1), and it is an essential step for our estimation of the phase error rate.

To obtain the yields $Y^{(Z)vir}_{sX,jX}$ we need to calculate

$$Y^{(Z)vir}_{sX,jX} = P_{Z_A}P_{Z_B}\,\mathrm{Tr}\left[\hat{D}_{sX}\hat{\theta}_{\mathrm{BE},jX,\mathrm{vir}}\right], \tag{13}$$

where $\hat{D}_{sX} = \sum \hat{A}^{\dagger}_k \hat{M}_{sX}\hat{A}_k$ corresponds to Eve's action, represented by the Kraus operators $\hat{A}_k$, as well as Bob's measurement with $\hat{M}_{sX}$ being an element of Bob's POVM. Here, recall the definition of the phase error rate where the Z basis is selected but both Alice and Bob use the X basis for their measurements (rather than the selected Z basis), which is why P_{Z_A} and P_{Z_B} appear in Eq. (13). Moreover, here we assume, for simplicity, that Eve applies the same quantum operation to every signal, which corresponds to a collective attack, but our analysis can be generalised to coherent attacks by considering the Azuma's inequality[33] (see section 'Azuma's inequality and its application to the security proof'), which deals with any correlations among the events, i.e., the phase error rate pattern. Using Eqs. (10)–(13), we obtain the following expression for the yields:

$$Y^{(Z)vir}_{sX,jX} = P_{Z_A}P_{Z_B}\Big(A_j\,\mathrm{Tr}\left[\hat{D}_{sX}|\gamma_{jX}\rangle\langle\gamma_{jX}|_{\mathrm{BE}}\right] + \mathrm{Tr}\left[\hat{D}_{sX}\Big(B_j|\gamma_{jX}\rangle\left\langle\gamma^{\perp}_{jX}\right|_{\mathrm{BE}} + B^*_j\left|\gamma^{\perp}_{jX}\right\rangle\langle\gamma_{jX}|_{\mathrm{BE}} + C_j\left|\gamma^{\perp}_{jX}\right\rangle\left\langle\gamma^{\perp}_{jX}\right|_{\mathrm{BE}}\Big)\right]\Big), \tag{14}$$

where the coefficients A_j, B_j and C_j are defined in section 'Coefficients', and we omit presenting their explicit expressions here for simplicity. As the state $|\gamma_{jX}\rangle_{\mathrm{BE}}$ in the first term of Eq. (14) is a single-mode qubit state, its density matrix can be expressed as

$$\hat{\rho}_{jX} = |\gamma_{jX}\rangle\langle\gamma_{jX}|_{\mathrm{BE}} = \frac{1}{2}\sum_i P^{jX,vir}_i\hat{\sigma}_i, \tag{15}$$

where $P^{jX,vir}_i$ are the coefficients of the Bloch vector and $\hat{\sigma}_i$, with $i \in \{Id, x, y, z\}$, represent the identity and the three Pauli operators, respectively. Therefore, we have that

$$A_j\,\mathrm{Tr}\left[\hat{D}_{sX}|\gamma_{jX}\rangle\langle\gamma_{jX}|_{\mathrm{BE}}\right] = A_j\left[P^{jX,vir}_{Id}q_{sX|Id} + P^{jX,vir}_x q_{sX|x} + P^{jX,vir}_y q_{sX|y} + P^{jX,vir}_z q_{sX|z}\right], \tag{16}$$

where $P^{jX,vir}_i = \mathrm{Tr}\left[\hat{\sigma}_i|\gamma_{jX}\rangle\langle\gamma_{jX}|_{\mathrm{BE}}\right]$ and $q_{sX|i} = \frac{1}{2}\,\mathrm{Tr}[\hat{D}_{sX}\hat{\sigma}_i]$ can be regarded as the transmission rates of the operator $\hat{\sigma}_i$. These can be calculated by solving a system of linear equations with the events from the actual protocol, which we will explain later. Moreover, by choosing the y-axis of the Bloch sphere appropriately we can always set $P^{jX,vir}_y = 0$ for all the Bloch vectors, as the PM just creates rotations in the X–Z plane of the Bloch sphere. Indeed, even if the PM introduces loss depending on Alice's state selection, as long as the three states form a triangle in the Bloch sphere, we can apply such simplification.[17] As already mentioned in section 'Assumptions on Alice's device', we note that any implementation of the LT protocol requires that the three states form a triangle in the Bloch sphere.

Furthermore, it is possible to find both lower and upper bounds on the second term of Eq. (14). In particular, this term can be written as $\mathrm{Tr}[\hat{D}_{sX}N_j]$ where N_j is the matrix $\begin{bmatrix} C_j & B^*_j \\ B_j & 0 \end{bmatrix}$ with eigenvalues

$$\lambda_{\max_j} = \frac{C_j + \sqrt{C_j^2 + 4|B_j|^2}}{2} \text{ and } \lambda_{\min_j} = \frac{C_j - \sqrt{C_j^2 + 4|B_j|^2}}{2}. \tag{17}$$

Using the properties of POVMs, we have that the operators $\hat{D}_{sX}$ have eigenvalues between 0 and 1; therefore, $\mathrm{Tr}[\hat{D}_{sX}N_j]$ is bounded by $\lambda_{\min_j} \le \mathrm{Tr}[\hat{D}_{sX}N_j] \le \lambda_{\max_j}$, since $\lambda_{\min_j}$ is negative.

This means that the virtual yields satisfy:

$$P_{Z_A}P_{Z_B}\left(A_j\left[q_{sX|Id} + P^{jX,vir}_x q_{sX|x} + P^{jX,vir}_z q_{sX|z}\right] + \lambda_{\min_j}\right) \le Y^{(Z)vir}_{sX,jX} \le P_{Z_A}P_{Z_B}\left(A_j\left[q_{sX|Id} + P^{jX,vir}_x q_{sX|x} + P^{jX,vir}_z q_{sX|z}\right] + \lambda_{\max_j}\right). \tag{18}$$

To find the transmission rates $q_{sX|i}$, the actual events we need to consider are those associated with the yields $Y^{(Z)}_{sX,0Z}$, $Y^{(Z)}_{sX,1Z}$ and $Y^{(X)}_{sX,0X}$. These are defined as $Y^{(\beta)}_{sX,j\beta} = P_{j\beta}P_{X_B}\,\mathrm{Tr}\left[\hat{D}_{sX}|\Phi_{j\beta}\rangle\langle\Phi_{j\beta}|_{\mathrm{BE}}\right]$ for $j\beta \in \{0Z, 1Z, 0X\}$, where the normalised actual states $|\Phi_{j\beta}\rangle_{\mathrm{BE}}$ are defined in Eq. (1). That is, $|\Phi_{j\beta}\rangle_{\mathrm{BE}}$ are the states emitted by Alice in the actual protocol when she chooses the bit value j and the basis β, in the presence of multi-mode signals. Using exactly the same method explained above, we obtain

$$Y^{(\beta)}_{sX,j\beta} = P_{j\beta}P_{X_B}\Big(E_{j\beta}\,\mathrm{Tr}\left[\hat{D}_{sX}|\phi_{j\beta}\rangle\langle\phi_{j\beta}|_{\mathrm{BE}}\right] + \mathrm{Tr}\left[\hat{D}_{sX}\left(F_{j\beta}|\phi_{j\beta}\rangle\left\langle\phi^{\perp}_{j\beta}\right|_{\mathrm{BE}} + F^*_{j\beta}\left|\phi^{\perp}_{j\beta}\right\rangle\langle\phi_{j\beta}|_{\mathrm{BE}} + G_{j\beta}\left|\phi^{\perp}_{j\beta}\right\rangle\left\langle\phi^{\perp}_{j\beta}\right|_{\mathrm{BE}}\right)\right]\Big), \tag{19}$$

where $E_{j\beta} = |a_{j\beta}|^2$, $F_{j\beta} = a_{j\beta}b^*_{j\beta}$, $F^*_{j\beta} = a^*_{j\beta}b_{j\beta}$ and $G_{j\beta} = |b_{j\beta}|^2$. Therefore, we find that the actual yields satisfy

$$P_{j\beta}P_{X_B}\left(E_{j\beta}\left[q_{sX|Id} + P^{j\beta}_x q_{sX|x} + P^{j\beta}_z q_{sX|z}\right] + \lambda_{\min_{j\beta}}\right) \le Y^{(\beta)}_{sX,j\beta} \le P_{j\beta}P_{X_B}\left(E_{j\beta}\left[q_{sX|Id} + P^{j\beta}_x q_{sX|x} + P^{j\beta}_z q_{sX|z}\right] + \lambda_{\max_{j\beta}}\right), \tag{20}$$

where

$$\lambda_{\max_{j\beta}} = \frac{G_{j\beta} + \sqrt{G^2_{j\beta} + 4|F_{j\beta}|^2}}{2} \text{ and } \lambda_{\min_{j\beta}} = \frac{G_{j\beta} - \sqrt{G^2_{j\beta} + 4|F_{j\beta}|^2}}{2}, \tag{21}$$

are the eigenvalues for the non-qubit part of the actual states, and $P^{j\beta}_x$ and $P^{j\beta}_z$ are the coefficients of the Bloch vector for the actual states. By substituting $j\beta \in \{0Z, 1Z, 0X\}$ in Eqs. (20) and (21), we obtain a system of three linear inequalities, which can be

expressed as

$$\begin{aligned}&\left[q_{sX|Id},\quad q_{sX|x},\quad q_{sX|z}\right]\hat{A}\\&+\left[\lambda_{\min_{0Z}},\quad \lambda_{\min_{1Z}},\quad \lambda_{\min_{0X}}\right]\le\left[\frac{Y^{(Z)}_{sX,0Z}}{P_{0Z}P_{X_B}},\quad \frac{Y^{(Z)}_{sX,1Z}}{P_{1Z}P_{X_B}},\quad \frac{Y^{(X)}_{sX,0X}}{P_{0X}P_{X_B}}\right]\\&\le\left[q_{sX|Id},\quad q_{sX|x},\quad q_{sX|z}\right]\hat{A}+\left[\lambda_{\max_{0Z}},\quad \lambda_{\max_{1Z}},\quad \lambda_{\max_{0X}}\right],\end{aligned}\tag{22}$$

where $\hat{A}:=(V^T_{0Z},V^T_{1Z},V^T_{0X})$ in which $V_{j\beta}=E_{j\beta}(1,P^{j\beta}_x,P^{j\beta}_z)$ and where the superscript T means transpose. By rearranging Eq. (22), we obtain the bounds on the transmission rates $q_{sX|Id}$, $q_{sX|x}$ and $q_{sX|z}$ to be

$$\begin{aligned}&\left(\left[\frac{Y^{(Z)}_{sX,0Z}}{P_{0Z}P_{X_B}},\quad \frac{Y^{(Z)}_{sX,1Z}}{P_{1Z}P_{X_B}},\quad \frac{Y^{(X)}_{sX,0X}}{P_{0X}P_{X_B}}\right]-\left[\lambda_{\max_{0Z}},\quad \lambda_{\max_{1Z}},\quad \lambda_{\max_{0X}}\right]\right)\hat{A}^{-1}\\&\le\left[q_{sX|Id},\quad q_{sX|x},\quad q_{sX|z}\right]\\&\le\left(\left[\frac{Y^{(Z)}_{sX,0Z}}{P_{0Z}P_{X_B}},\quad \frac{Y^{(Z)}_{sX,1Z}}{P_{1Z}P_{X_B}},\quad \frac{Y^{(X)}_{sX,0X}}{P_{0X}P_{X_B}}\right]-\left[\lambda_{\min_{0Z}},\quad \lambda_{\min_{1Z}},\quad \lambda_{\min_{0X}}\right]\right)\hat{A}^{-1},\end{aligned}\tag{23}$$

where $\hat{A}^{-1}$ is the inverse of the matrix $\hat{A}$.

By solving Eq. (23), we can calculate the transmission rates and then substitute them into Eq. (18) to find the upper bounds on the virtual yields $Y^{(Z)vir}_{0X,1X}$ and $Y^{(Z)vir}_{1X,0X}$. Finally, by using these upper bounds on the virtual yields and the yields from the actual events we can estimate the phase error rate e_X in Eq. (7).

As already mentioned previously, this technique is quite general and could be applied to many other QKD protocols. As an example, in the Supplementary Material (Section II), we outline how this analysis could be performed for MDI-QKD.

Simulation of the key rate

Only for the purpose of the simulation, we now consider a particular device model and a particular THA. In general, to experimentally guarantee that the three states emitted by Alice remain in two dimensions, i.e., in a single-mode qubit, her PM needs to have the same temporal, spectral, spatial and polarisation mode independently of the bit and basis choices. However, due to imperfections in the devices this condition is hard to fulfil. Some counter-measures against these imperfections have been suggested,[34–37] but they cannot rigorously ensure a single-mode qubit. Therefore, it is crucial to consider how device's flaws can be taken into account in a security proof. This is the aim of our analysis. For simplicity, among many imperfections, we select the polarisation mode as an example of how to use our framework.

A change in polarisation can arise from the imperfect alignment of the laser with the principal axis of the PM and/or when the PM is polarisation dependent, i.e., the state of polarisation of the signals prepared might be different for each encoding phase value. In principle, this could be avoided by using a polarisation beamsplitter (PBS) that selects a single polarisation mode. In practice, however, because of the finite extinction ratio of the PBS this is usually not the case. Here we relax the need for a perfect PBS by considering a polarisation multi-mode scenario. We remark, nonetheless, that our analysis can be applied to any multi-mode scenario. Using our formalism, we can express the states sent by Alice in the scenario considered in an analogous way to Eq. (1):

$$\left|\Omega_{j\beta}\right\rangle_B=\cos\theta_{j\beta}\left|\omega_{j\beta}\right\rangle_{HB}+\sin\theta_{j\beta}\left|\omega_{j\beta}\right\rangle_{VB},\tag{24}$$

for $j\beta\in\{0Z,\,1Z,\,0X\}$, where the subscripts H and V refer to the horizontal and vertical polarisation modes, respectively. That is, now the polarisation state of $|\omega_{j\beta}\rangle_B$ depends on Alice's bit and basis choices instead of being the same independently of her encoding. Next, we add the SPF and the THA to this particular device model.

For the states $|\omega_{j\beta}\rangle_{HB}$ and $|\omega_{j\beta}\rangle_{VB}$, we use the definitions in Eq. (3), where they both live in a qubit space. Also, by using Eq. (3) these states already include SPFs whenever the parameter $\delta>0$. As stressed in section 'Assumptions on Alice's device', as in this case we know the form of the states we do not need to consider the worst-case scenario but only the inner product ${}_{HB}\langle\omega_{j\beta}|\omega_{j'\beta'}\rangle_{VB}=0$ for all j, j', β and β'.

In addition, we consider an active information leakage in our device model. For this, we assume that Eve sends strong light into Alice's PM, which is then back-reflected and exits Alice's lab in the form

$$\left|\xi_{j\beta}\right\rangle_E=C_I\left|e\right\rangle_E+C_D\left|e_{j\beta}\right\rangle_E.\tag{25}$$

In this expression, $|C_I|^2+|C_D|^2=1$ and $|e\rangle_E$ ($|e_{j\beta}\rangle_E$) represents (represent) the setting independent (dependent) state (states) on Alice's bit and basis choice, where we assume that $\langle e|e_{j\beta}\rangle_E=0$. That is, the state $|e\rangle_E$ ($|e_{j\beta}\rangle_E$) provides Eve with no (some) information about Alice's bit and basis values each given time. Therefore, our model for the THA can be parameterised by only two parameters, C_I and C_D, and no further detailed information is needed to apply our analysis. For instance, when we increase isolation on Alice's sending device, the independent component increases and Eve obtains less information about the states being sent. Moreover, in the absence of further information about the states $|e_{j\beta}\rangle_E$, we assume the worst-case scenario where these states are orthogonal to each other, i.e., $\langle e_{j\beta}|e_{j'\beta}\rangle_E=0$ for any $(j,\beta)\neq(j',\beta')$. Clearly, if Alice and Bob know the states $|e_{j\beta}\rangle_E$, this information can be trivially included in the formalism below.

If $|\xi_{j\beta}\rangle_E$ is say, for instance, a coherent state, $|e\rangle_E$ is the vacuum state (i.e., $|e\rangle_E=|v\rangle_E$), $C_I=e^{-\mu/2}$ and $C_D=\sqrt{1-e^{-\mu}}$, where μ is the intensity of Eve's back-reflected light. In this case, note that the condition $\langle e_{j\beta}|e_{j'\beta'}\rangle_E=0$ is not satisfied, as Eve will never be able to perfectly distinguish the states dependent on Alice's encoding. The value of this overlap depends on the isolation of the devices. Below, however, we conservatively assume for simplicity the worst-case scenario where this overlap is zero.

Putting Eqs. (24) and (25) together, Alice's emitted state for the single-photon pulses is modelled as

$$\left|\Phi_{j\beta}\right\rangle_{BE}=\left|\Omega_{j\beta}\right\rangle_B\otimes\left|\xi_{j\beta}\right\rangle_E.\tag{26}$$

By using Eqs. (3), (24), (25) and (26), and by assuming that $|\xi_{j\beta}\rangle_E$ are coherent states, we obtain

$$\begin{aligned}\left|\Phi_{j\beta}\right\rangle_{BE}&=\left(\cos\theta_{j\beta}\left|\omega_{j\beta}\right\rangle_{HB}+\sin\theta_{j\beta}\left|\omega_{j\beta}\right\rangle_{VB}\right)\otimes\left(C_I\left|v\right\rangle_E+C_D\left|e_{j\beta}\right\rangle_E\right)\\&=\cos\theta_{j\beta}C_I\left|\omega_{j\beta}\right\rangle_{HB}\left|v\right\rangle_E+\cos\theta_{j\beta}C_D\left|\omega_{j\beta}\right\rangle_{HB}\left|e_{j\beta}\right\rangle_E\\&\quad+\sin\theta_{j\beta}\left|\omega_{j\beta}\right\rangle_{VB}\otimes\left(C_I\left|v\right\rangle_E+C_D\left|e_{j\beta}\right\rangle_E\right).\end{aligned}\tag{27}$$

The first term of Eq. (27) has polarisation H and is insensitive to the THA, it corresponds to $a_{j\beta}|\phi_{j\beta}\rangle_{BE}$ in Eq. (1). Similarly, the other terms have either polarisation V and/or are affected by the THA, and together they correspond to $b_{j\beta}\left|\phi^{\perp}_{j\beta}\right\rangle_{BE}$ in Eq. (1). In this case, the unnormalised virtual states given by Eq. (10) have now the form

$$\begin{aligned}\left|\psi\right\rangle_{BE,jX,vir}=\frac{1}{2}\Bigg[&C_I\sqrt{\cos\theta^2_{0Z}-(-1)^j2\cos\theta_{0Z}\cos\theta_{1Z}\sin\tfrac{\delta}{2}+\cos\theta^2_{1Z}}\left|\gamma_{jX}\right\rangle_{BE}\\&+\sqrt{C_I^2\left(\sin\theta^2_{0Z}-(-1)^j2\sin\theta_{0Z}\sin\theta_{1Z}\sin\tfrac{\delta}{2}+\sin\theta^2_{1Z}\right)+2C_D^2}\left|\gamma^{\perp}_{jX}\right\rangle_{BE}\Bigg],\end{aligned}\tag{28}$$

where we have used the relationship $\langle\omega_{0Z}|\omega_{1Z}\rangle_B=\langle\omega_{0Z}|\omega_{1Z}\rangle_B=-\sin(\frac{\delta}{2})$. In order to estimate the phase error rate, we need to calculate the transmission rates $q_{sX|Id}$, $q_{sX|x}$ and $q_{sX|z}$ using the actual yields. For this, we use Eq. (23) where in

this particular example, the matrix $\hat{A}$ is

$$\hat{A} = \begin{bmatrix} E_{0Z} & E_{1Z} & E_{0X} \\ 0 & -E_{1Z}\sin(\delta) & E_{0X}\sin(\pi/2+\delta/2) \\ E_{0Z} & -E_{1Z}\cos(\delta) & E_{0X}\cos(\pi/2+\delta/2) \end{bmatrix}, \tag{29}$$

where $E_{j\beta} = C_l^2 \cos^2\theta_{j\beta}$. Then, we can find the virtual yields by using Eq. (18) where, in this example, the coefficients of the Bloch vectors are

$$\begin{aligned} P_x^{jX,vir} &= \frac{(-1)^j 2\cos\theta_{0Z}\cos\theta_{1Z}\cos\frac{\delta}{2} - 2\cos\theta_{1Z}^2\cos\frac{\delta}{2}\sin\frac{\delta}{2}}{\cos\theta_{0Z}^2 - (-1)^j 2\cos\theta_{0Z}\cos\theta_{1Z}\sin\frac{\delta}{2} + \cos\theta_{1Z}^2}, \\ P_z^{jX,vir} &= \frac{\cos\theta_{0Z}^2 - (-1)^j 2\cos\theta_{0Z}\cos\theta_{1Z}\sin\frac{\delta}{2} + \cos\theta_{1Z}^2(1-2\cos^2\frac{\delta}{2})}{\cos\theta_{0Z}^2 - (-1)^j 2\cos\theta_{0Z}\cos\theta_{1Z}\sin\frac{\delta}{2} + \cos\theta_{1Z}^2}. \end{aligned} \tag{30}$$

Finally, one can directly use Eq. (7) to estimate the phase error rate e_X.

Results of the simulation

With the method described above, it is possible to employ the leaky source without compromising the security of the QKD system. Nonetheless, depending on the situation and the particular experimental parameters, it might be beneficial to consider another method that, in some cases, might provide a higher key generation rate. Therefore, it is important to compare our generalised LT protocol with an alternative method, say the LP analysis introduced in ref. [29]. In Section III of Supplementary Material, we provide a detailed description of this analysis.

In order to evaluate how the different imperfections of the source affect the key generation rate for both security proofs, we analyse each of them separately. The results and discussion of this analysis are in the Supplementary Material (Sections I and III). Now, we present the comparison results between the generalised LT protocol and the LP analysis, and identify which one provides a better R depending on the experimental setup. This way an experimentalist can choose which method to use for known device parameters and ensure the security of the generated key between Alice and Bob. It is noteworthy that for a fair comparison between both protocols we consider the efficient four-state LT protocol, where the four states of the BB84 protocol[38] are used to run two LT protocols simultaneously. That is, when Alice emits the state $|\omega_{0X}\rangle_B$ ($|\omega_{1X}\rangle_B$), she considers that it belongs to the first (second) LT protocol, whereas each of the two protocols is randomly chosen by Alice before sending the pulse. See Supplementary Eq. (III.8) for the definition of the state $|\omega_{1X}\rangle_B$ in the presence of SPFs. This means that no modifications to the hardware of the standard BB84 protocol are required and therefore, the four-state LT protocol is equivalent to the BB84 protocol from an experimental point of view.

We show the results obtained for R as a function of the overall system loss (which includes both the channel attenuation and the loss at Bob's receiver), for different values of δ, $\theta_{j\beta}$ and μ, which correspond to the SPFs, non-qubit assumption and THA, respectively. The angles $\theta_{j\beta}$ are chosen such that they are associated with Alice's encoding of the states $|\omega_{j\beta}\rangle_B$. That is, $\theta_{0Z} = 0$, $\theta_{1Z} = \pi\hat{\theta}$ and $\theta_{0X} = \frac{\pi}{2}\hat{\theta}$ for a certain angle $\hat{\theta}$. In our simulations, we consider the

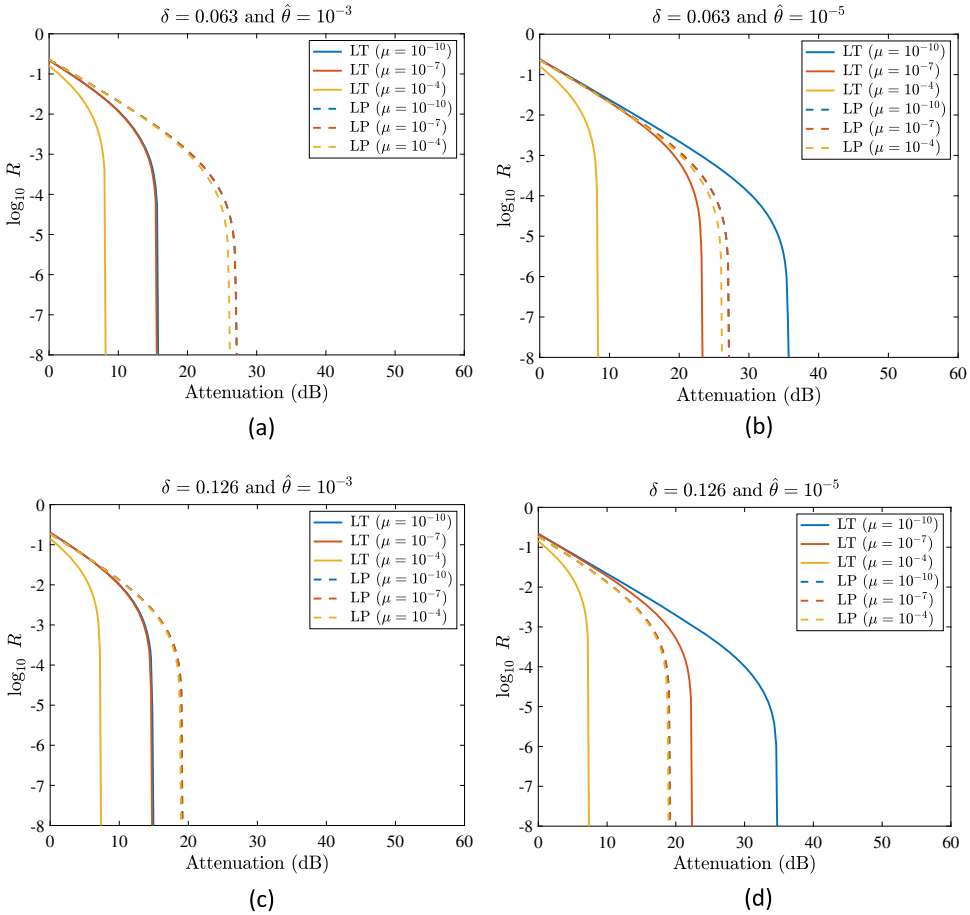

Fig. 2 Secret key rate R vs. the overall system loss measured in dB for the generalised LT protocol and LP analysis. The blue and red dashed lines are superimposed in all graphs. **a** The LP protocol performs better in this scenario, because the SPFs are small but $\hat{\theta}$ is high. **b** For a smaller $\hat{\theta}$, the generalised LT analysis is better when $\mu = 10^{-10}$. **c** The LP performs better when $\hat{\theta}$ is larger even if δ is high. **d** For large δ and small $\hat{\theta}$, the generalised LT clearly surpasses the LP analysis when $\mu = 10^{-10}$ or $\mu = 10^{-7}$

experimental parameters to be the following: the dark count rate $p_d = 10^{-7}$, $f = 1.16$ and the fiber loss coefficient $\alpha = 0.2$ dB/Km. Moreover, we assume for simplicity that in the LT protocol and in the LP analysis $P_{Z_A} = \frac{1}{2}$ and $P_{Z_B} = \frac{1}{2}$. This selection of probabilities might not be ideal but it is sufficient for the purpose of the simulation. By using the channel model described in the Supplementary Material (Section IV), we find that $Y_Z = Y^{(Z)}_{0Z,0Z} + Y^{(Z)}_{1Z,0Z} + Y^{(Z)}_{0Z,1Z} + Y^{(Z)}_{1Z,1Z} = P_{Z_A}P_{Z_B}[4(1-\frac{\eta}{2})p_d + \eta]$ where η is the overall transmission efficiency of the system (see Section IV in the Supplementary Material for more details). The bit error rate is then given by

$$e_Z = \frac{2(1-\frac{\eta}{2})p_d + \frac{\eta}{2} + \frac{\eta}{4}(\cos 2\delta + \cos\delta)(p_d - 1)}{4(1-\frac{\eta}{2})p_d + \eta}. \quad (31)$$

As the range of the experimental parameters greatly depend on the devices being used, we select the SPFs to be either $\delta = 0.063$ or $\delta = 0.126$ according to the experimental results reported in refs. [34,39,40]. There are some works related with the mode dependency,[34,41] but unfortunately they do not directly provide the value of $\hat{\theta}$. Therefore, we evaluate $\hat{\theta}$ over a big range (see Sections I and III in the Supplementary Material) and for these simulations we choose $\hat{\theta} = 10^{-3}$ and $\hat{\theta} = 10^{-5}$. Obviously, a better experimental characterisation of the source would be essential to improve the accuracy of the current parameters. Finally, for the intensity of Eve's back-reflected light during the THA, we use $\mu = 10^{-10}$, $\mu = 10^{-7}$ and $\mu = 10^{-4}$.[21] The results are shown in Fig. 2. It is noteworthy that the blue and red dashed lines coincide (for the resolution presented) in all graphs. The reason for this lies in the value of the variable μ. That is, for $\mu = 10^{-10}$ and $\mu = 10^{-7}$, the LP analysis results in approximately the same secret key rate (see Supplementary Material, Section III, for more details).

By comparing Fig. 2a, c or Fig. 2b, d, we can see how an increase in the parameter δ, which is associated with SPFs, affects both protocols. For the generalised LT protocol, the key rate stays approximately the same as expected, as this method is loss tolerant to SPFs. On the other hand, the LP analysis is more influenced by SPFs (see also Supplementary Fig. III.2). The reason for this difference is that in LP it is assumed the worst-case scenario, in which Eve can enhance the basis dependence of the signals by exploiting the channel loss. However, in LT no such assumption is required; hence, the performance is maintained. This means that LT will typically outperform LP in the presence of high SPFs.

To compare LT and LP as a function of the setting-dependent $\theta_{j\beta}$, we can contrast Fig. 2a, b or Fig. 2c, d. The graphs show clear differences due to decreasing the value of $\hat{\theta}$, especially for the LT case. In Fig. 2a, LP reaches a longer distance for any value of μ, but when $\hat{\theta} = 10^{-5}$ LT gets better, particularly for $\mu = 10^{-10}$ as seen in Fig. 2b. Furthermore, Fig. 2b shows that even when there are SPFs, LP can still do better than LT if the states sent are far from the idealised qubit. This is because the non-qubit assumption negatively affects more LT than LP (see Supplementary Figs I.1c and III.2c).

When we compare the values of μ for the LT and LP, we can see a similar trend in the secret key rate for all graphs in Fig. 2. Namely, the difference between the curves when $\mu = 10^{-10}$ (blue) and $\mu = 10^{-4}$ (yellow) is much larger for LT than for LP, which means that the THA is worse for the LT. However, μ is a parameter that might be easily controlled experimentally by introducing passive counter-measures, such as optical isolators.[21] Indeed, in ref. [21], it has been shown, for instance, that a value of $\mu = 10^{-6}$ could be easily achieved in practice. For example, even if Eve sends Alice optical pulses with 10^{20} photons, practical combinations of the components of Alice's transmitter could guarantee a total optical isolation of -170 dB, which would be enough to achieve $\mu = 10^{-6}$.[21] This means that the LT method may be a better alternative when the SPFs are more dominant and the mode dependency is small, as it outperforms the LP analysis in Fig. 2b and d.

As explained above, the non-qubit assumption and the THA affect more the LT than the LP analysis. This might be because our generalisation of the LT protocol is overestimating Eve. When we calculate the bounds for the yields we obtain that the eigenvalues λ_{max} and λ_{min} depend on the state preparation. However, this is probably too pessimistic, because there might be some additional constraints among them, as the space spanned by the states associated to $0Z$ and $1Z$, respectively, is not orthogonal to the one spanned by the virtual states associated to $0X^{vir}$ and $1X^{vir}$. This means that these separate optimisations should not be possible in practice, because Eve cannot achieve optimal values for all λs. In other words, by improving our characterisation of the states we can improve the performance of the generalised LT protocol. This is however beyond the scope of this paper and we leave it for future work.

In order to further investigate the differences between the two methods, we determine the parameter regimes where their performance is identical. First, by setting $\hat{\theta} = 10^{-6}$, we can identify which values of δ and μ provide the same key generation rate R for LT and LP. The results are presented in Fig. 3a, where the diagram clearly shows which protocol performs better given a certain δ and μ: above the fitted curve, the LT provides a better performance but below the curve LP is the preferable method. In other words, as the SPFs increase the LT is superior but as μ increases LP becomes more suitable.

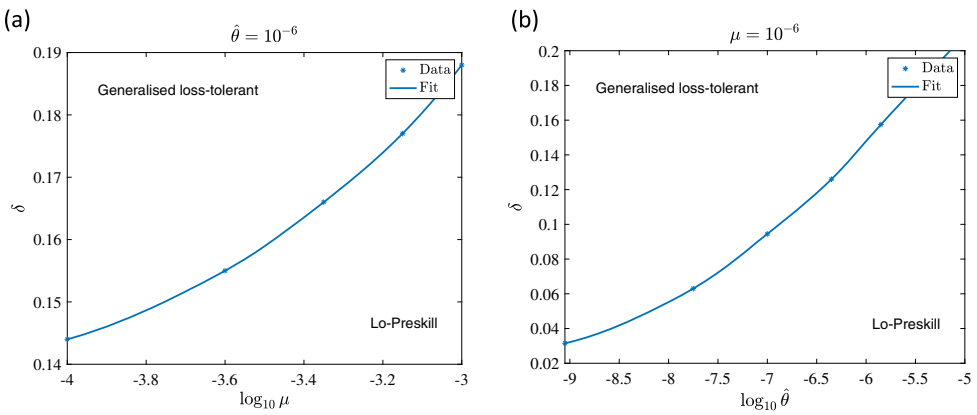

Fig. 3 The fitted line corresponds to those experimental parameters that result in the same key generation rate R for both methods, the generalised LT protocol and the LP analysis. Above the line the generalised LT protocol performs better, and below the line the LP analysis is the preferred method. The data points were fitted using a shape-preserving interpolant in Matlab. **a** Plot of δ against μ for $\hat{\theta} = 10^{-6}$. **b** Plot of δ against $\hat{\theta}$ for $\mu = 10^{-6}$

Similar results are obtained when $\mu = 10^{-6}$. This case is particularly useful, as in principle we can control the value of μ experimentally by the amount of isolation we use in our devices. Again, as SPFs increase the LT becomes better, giving a better estimation of the phase error rate and a better secret key generation rate.

DISCUSSION

Typical security proofs ignore many imperfections of experimental devices, thus hindering the security claim of QKD. In this work, we have generalised the LT QKD protocol to accommodate general imperfections. In particular, our formalism is valid for a general device model with, for instance, SPFs, mode dependency and THAs, which result in passive and active information leakage to an eavesdropper. Using this multi-mode scenario, we have shown that the qubit assumption can be removed from the LT protocol without compromising the security of the QKD scheme. We present a formalism that can be used to estimate the phase error rate by finding the transmission rates of some virtual states and assuming the general state structure defined in Eq. (1). Therefore, in principle, it can be applied to most QKD protocols.

In order to compare our generalised LT protocol with other security proofs, we have applied the LP analysis[29] to the same device model. In so doing, we have identified which approach delivers a higher secret key rate as a function of the experimental parameters. For example, the results obtained show that LP method performs better under the non-qubit assumption and the THA but the generalised LT protocol is better when there are SPFs. As the THAs can be controlled using passive counter-measures, such as optical isolators, we have shown that in some cases the generalised LT protocol might be the preferable method when the SPFs are more dominant. This way, our work can be used as a guideline to improve current experimental implementations in which multi-mode QKD is unavoidable. Moreover, it highlights the importance of source characterisation for more realistic security proofs.

For completeness, we also note that ref. [42] has recently proposed a computational toolbox that can be used to numerically estimate the phase error rate of a QKD protocol and such technique could be applied to the scenario considered in this paper. Essentially, similar to the LP analysis, their technique only requires the knowledge of the inner products between the states emitted by Alice and is mathematically simple, which is a striking difference to previous numerical analyses.[43,44] That is, the approach in ref. [42] can also remove the qubit assumption and include side channels when estimating the phase error rate. There are, however, some relevant differences between that method and our formalism, besides the obvious one, i.e., that our work is an analytical technique. The approach in ref. [42] requires a full characterisation of the side channels in order to obtain the inner product of the states, whereas ours does not, resulting in a simpler characterisation of the source. Moreover, in the absence of side channels, their method is not loss tolerant in some parameter regimes, whereas ours is always loss tolerant, which is essential to guarantee a good performance over long distances. Furthermore, their analysis considers pure states, whereas our method also applies to the mixed-state scenario. Despite these differences, it would be interesting to combine the advantages of both methods to achieve a better implementation security, but we leave this for future works.

METHODS

Security proof against coherent attacks

Here we present the security proof of our formalism against coherent attacks. For simplicity of the discussion, the Results section deals with the case of pure states in a single-mode qubit space; however, in this section we consider the general scenario where the states could be mixed states in a single-mode qubit space. For this, we consider a virtual protocol.[45,46] This protocol is equivalent to the actual protocol in the sense that the resulting statistics of the measurements and the secret key rate generated between Alice and Bob are the same. Furthermore, the classical and quantum information available to Eve is equal in both protocols. The security claim follows from the fact that Alice and Bob can choose which protocol to execute and Eve is unable to distinguish between them. Hence, by proving the security of the virtual protocol we prove the security of the actual protocol.

In this work we employ the complementary scenario,[31,32] which considers a virtual protocol that uses the complementary observable of the key generation basis. For instance, in the actual protocol Alice and Bob agree on the bit values in the Z basis, whereas in the virtual protocol they collaborate to prepare a qubit in an eigenstate of the X basis. In doing so, the security proof basically reduces to the estimation of the phase error rate, which corresponds to the bit error rate that Alice and Bob would have observed if they would have measured the Z basis state in the X basis. Therefore, the aim of the virtual protocol is to estimate the phase error rate. In the section 'Estimation of the phase error rate', we showed how this can be done by using our formalism and how we can calculate the secret key rate R against collective attacks. Here we describe in detail the virtual protocol used for the security proof and explain how to accommodate coherent attacks by Eve through the use of Azuma's inequality.[33]

We consider a more general case than that studied in the Results section in which Alice generates a single-mode qubit system B, whose states are mixed states, and we show how to define the pure states needed for our security proof. We denote the mixed states by the density matrices $\hat{\rho}_{0Z_B}$, $\hat{\rho}_{1Z_B}$ and $\hat{\rho}_{0X_B}$. These states are diagonalised as

$$\begin{aligned}\hat{\rho}_{jZ_B} &= P^0_{jZ}\left|\phi^0_{jZ}\right\rangle\left\langle\phi^0_{jZ}\right|_B + P^1_{jZ}\left|\phi^1_{jZ}\right\rangle\left\langle\phi^1_{jZ}\right|_B,\\ \hat{\rho}_{0X_B} &= P^0_{0X}\left|\phi^0_{0X}\right\rangle\left\langle\phi^0_{0X}\right|_B + P^1_{jZ}\left|\phi^1_{0X}\right\rangle\left\langle\phi^1_{0X}\right|_B,\end{aligned} \tag{32}$$

where $j \in \{0, 1\}$ and P^0_{jZ}, P^1_{jZ}, P^0_{0X} and P^1_{0X} are probabilities satisfying $P^0_{jZ} + P^1_{jZ} = 1$ and $P^0_{0X} + P^1_{0X} = 1$. Moreover, $\left\{\left|\phi^0_{jZ}\right\rangle_B, \left|\phi^1_{jZ}\right\rangle_B\right\}$ and $\left\{\left|\phi^0_{0X}\right\rangle_B, \left|\phi^1_{0X}\right\rangle_B\right\}$ are orthonormal bases in the single-mode qubit. The states sent might be mixed due to imperfections in Alice's devices, including a potential entanglement between her devices and Eve's ancilla. This means that in general these mixed states can be purified by introducing Alice's ancilla system A_1 and Eve's system E, and therefore we have the purifications of $\hat{\rho}_{0Z_B}$, $\hat{\rho}_{1Z_B}$ and $\hat{\rho}_{0X_B}$ as $\left|\tilde{\psi}_{0Z}\right\rangle_{A_1BE}$, $\left|\tilde{\psi}_{1Z}\right\rangle_{A_1BE}$, and $\left|\tilde{\psi}_{0X}\right\rangle_{A_1BE}$, each of which expressed by

$$\begin{aligned}\left|\tilde{\psi}_{jZ}\right\rangle_{A_1BE} &= \sqrt{P^0_{jZ}}\left|0_{jZ}\right\rangle_{A_1E}\left|\phi^0_{jZ}\right\rangle_B + \sqrt{P^1_{jZ}}\left|1_{jZ}\right\rangle_{A_1E}\left|\phi^1_{jZ}\right\rangle_B,\\ \left|\tilde{\psi}_{0X}\right\rangle_{A_1BE} &= \sqrt{P^0_{0X}}\left|0_{0X}\right\rangle_{A_1E}\left|\phi^0_{0X}\right\rangle_B + \sqrt{P^1_{0X}}\left|1_{0X}\right\rangle_{A_1E}\left|\phi^1_{0X}\right\rangle_B.\end{aligned} \tag{33}$$

Here, $\left\{\left|0_{jZ}\right\rangle_{A_1E}, \left|1_{jZ}\right\rangle_{A_1E}\right\}$ and $\left\{\left|0_{0X}\right\rangle_{A_1E}, \left|1_{0X}\right\rangle_{A_1E}\right\}$ are orthonormal bases. Now, we define states similar to Eq. (5) that include the purification of Alice's state:

$$\begin{aligned}\left|\tilde{\Psi}_Z\right\rangle_{A_1A_2BE} &= \tfrac{1}{\sqrt{2}}\left[\left|0_Z\right\rangle_{A_2}\left|\tilde{\psi}_{0Z}\right\rangle_{A_1BE} + \left|1_Z\right\rangle_{A_2}\left|\tilde{\psi}_{1Z}\right\rangle_{A_1BE}\right],\\ \left|\tilde{\Psi}_X\right\rangle_{A_1A_2BE} &= \left|0_X\right\rangle_{A_2}\left|\tilde{\psi}_{0X}\right\rangle_{A_1BE},\end{aligned} \tag{34}$$

where A_2 is Alice's ancilla system used to generate a bit value in the protocol, i.e., it possesses information about Alice's encoding. As explained above, in the security analysis Alice measures A_2 in the X basis instead of the Z basis when $\left|\tilde{\Psi}_Z\right\rangle_{A_1A_2BE}$ is prepared; therefore, it is useful to write this state in the X basis of system A_2. By substituting $\left|0_Z\right\rangle_{A_2} = \frac{1}{\sqrt{2}}\left(\left|0_X\right\rangle_{A_2} + \left|1_X\right\rangle_{A_2}\right)$ and $\left|1_Z\right\rangle_{A_2} = \frac{1}{\sqrt{2}}\left(\left|0_X\right\rangle_{A_2} - \left|1_X\right\rangle_{A_2}\right)$ we can express $\left|\tilde{\Psi}_Z\right\rangle_{A_1A_2BE}$ as

$$\begin{aligned}\left|\tilde{\Psi}_Z\right\rangle_{A_1A_2BE} = &\sqrt{\frac{1+\left\langle\tilde{\psi}_{0Z}|\tilde{\psi}_{1Z}\right\rangle_{A_1BE}}{2}}\left|0_X\right\rangle_{A_2}\left|\tilde{\psi}^{vir}_{0X}\right\rangle_{A_1BE}\\ &+ \sqrt{\frac{1-\left\langle\tilde{\psi}_{0Z}|\tilde{\psi}_{1Z}\right\rangle_{A_1BE}}{2}}\left|1_X\right\rangle_{A_2}\left|\tilde{\psi}^{vir}_{1X}\right\rangle_{A_1BE},\end{aligned} \tag{35}$$

where

$$\left|\tilde{\psi}_{jX}^{vir}\right\rangle_{A_1BE}=\frac{\left|\tilde{\psi}_{0Z}\right\rangle_{A_1BE}+(-1)^j\left|\tilde{\psi}_{1Z}\right\rangle_{A_1BE}}{\sqrt{2\left(1+(-1)^j\left\langle\tilde{\psi}_{0Z}|\tilde{\psi}_{1Z}\right\rangle_{A_1BE}\right)}}. \tag{36}$$

In the virtual protocol, we consider that Alice sends Bob two virtual states, $\left|\tilde{\psi}_{jX}^{vir}\right\rangle_{A_1BE}$, and three actual states, $\left|\tilde{\psi}_{jZ}\right\rangle_{A_1BE}$ and $\left|\tilde{\psi}_{0X}\right\rangle_{A_1BE}$, which are used to estimate the phase error rate. We have seen that even in the case of mixed states we can define actual and virtual pure states, and these pure states can be directly used in our security proof. Therefore, our formalism is valid for mixed states in a single-mode qubit space.

Next, let us continue to explain the security proof in more detail. The selection of these actual and virtual states can be expressed as

$$|\varphi\rangle_{SA_1BE}=\sum_{c=1}^{5}\sqrt{P(c)}|c\rangle_S\left|\vartheta^{(c)}\right\rangle_{A_1BE}, \tag{37}$$

where S is the shield system that is kept inside of Alice's lab and the states $\left|\vartheta^{(c)}\right\rangle_{A_1BE}$ are

$$\begin{aligned}
\left|\vartheta^{(1)}\right\rangle_{A_1BE}&=\left|\tilde{\psi}_{0X}^{vir}\right\rangle_{A_1BE},\\
\left|\vartheta^{(2)}\right\rangle_{A_1BE}&=\left|\tilde{\psi}_{1X}^{vir}\right\rangle_{A_1BE},\\
\left|\vartheta^{(3)}\right\rangle_{A_1BE}&=\left|\tilde{\psi}_{0Z}\right\rangle_{A_1BE},\\
\left|\vartheta^{(4)}\right\rangle_{A_1BE}&=\left|\tilde{\psi}_{1Z}\right\rangle_{A_1BE},\\
\left|\vartheta^{(5)}\right\rangle_{A_1BE}&=\left|\tilde{\psi}_{0X}\right\rangle_{A_1BE},
\end{aligned} \tag{38}$$

with their respective probabilities $P(c)$

$$\begin{aligned}
P(1)&=\tfrac{P_{Z_A}P_{Z_B}}{2}\left(1+\left\langle\tilde{\psi}_{0Z}|\tilde{\psi}_{1Z}\right\rangle_{A_1BE}\right),\\
P(2)&=\tfrac{P_{Z_A}P_{Z_B}}{2}\left(1-\left\langle\tilde{\psi}_{0Z}|\tilde{\psi}_{1Z}\right\rangle_{A_1BE}\right),\\
P(3)&=\tfrac{P_{Z_A}P_{X_B}}{2},\\
P(4)&=\tfrac{P_{Z_A}P_{X_B}}{2},\\
P(5)&=P_{X_A}P_{Z_B}+P_{X_A}P_{X_B}=P_{X_A}.
\end{aligned} \tag{39}$$

When Bob receives the states, he performs a measurement in either the Z or the X basis, and these are defined by the POVMs described in section 'Assumptions on Bob's device'. Also, all announcements between Alice and Bob are done via an authenticated public channel. It is noteworthy that in the virtual protocol we assume that Alice and Bob are sitting in the same lab so that they can choose the measurement basis, and this is allowed because the quantum and classical information available to Eve is the same between the actual and the virtual protocols. The detailed steps of the virtual protocol are presented below and the logic schematics in Fig. 4.

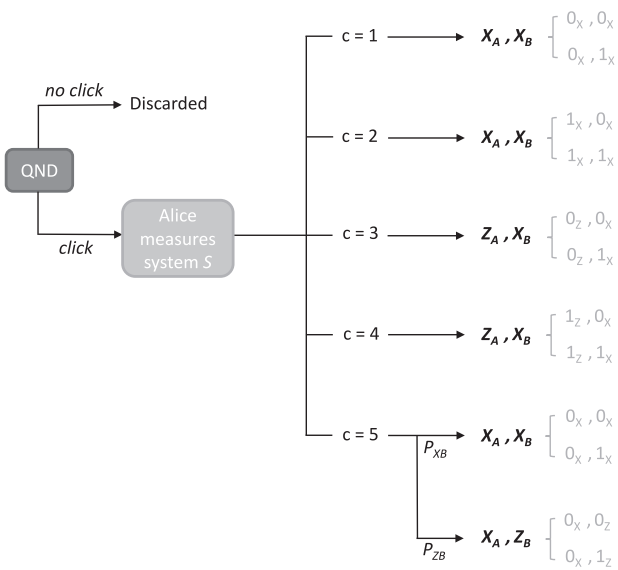

Fig. 4 The logical schematics for the virtual protocol, where the notation X_A/Z_A, X_B/Z_B corresponds to Alice's and Bob's measurements bases, respectively. The virtual states correspond to $c=1$, 2, the actual Z states to $c=3$, 4, and the actual X states to $c=5$. For each click event, Alice measures system S and Bob measures system BE. It is noteworthy that the selection of $c=1$, 2, 3, 4 already includes Bob's measurement in the X basis, but when $c=5$ his measurement basis is chosen probabilistically

Virtual protocol

1. **Initialisation:** Before running the protocol, Alice and Bob agree on a number N_{fixed} of rounds, on the error correcting codes and on a set of hash functions to perform privacy amplification. Steps 2–4 of the protocol are repeated N times until the number of detected events N becomes N_{fixed}.
2. **State preparation:** After a potential THA, Alice prepares systems S, A_1 and BE in the entangled state $|\varphi\rangle_{SA_1BE}$, in Eq. (37), and sends Bob the system BE via a quantum channel.
3. **QND measurement:** For each incoming system, Bob performs a quantum non-demolition (QND) measurement to determine whether the signals are detected or not. If Bob obtains a detection event, he keeps the resulting system and N is increased by 1 unit.
4. **Detection announcement:** If $N=N_{\text{fixed}}$, Bob announces the termination of quantum communication and the detection pattern. Otherwise Alice and Bob return to Step 2 of the protocol.
5. **Measurement and basis announcement:** For each of the detected events, Alice measures her system S and announces the Z (X) basis when $c=1$, 2, 3, 4 ($c=5$). Bob announces the Z (X) basis for $c=1$, 2 (3, 4), but he always measures in the X basis. For $c=5$, Bob selects the basis $\beta\in\{Z, X\}$ probabilistically and announces his basis choice. Then, he carries out the measurement on system BE in his selected basis.
6. **Sifting and announcement:** Alice and Bob define and announce the bit strings $\vec{s}_{X,0Z}$, $\vec{s}_{X,1Z}$ and $\vec{s}_{X,0X}$, which correspond to the events when Alice sends the actual states and Bob performs the X basis measurements. These are the basis mismatched events when $c=3$, 4 and one of the events when $c=5$, the basis matched event. These strings are used to estimate the phase error rate.

In the virtual protocol, we require that Alice and Bob postpone their measurements until the quantum communication ends; therefore, we assume that Alice and Bob possess quantum memories where they can store their systems. The reason for this deferral comes from the application of Azuma's inequality, which is explained later. In the case of Alice, she only makes her measurement after the termination condition, in Step 5. This is allowed because it does not matter when she performs the measurement, as it commutes with Eve's operations and hence it will not affect Alice's statistics. For Bob, we divide his measurement in two steps: a QND measurement, which allows him to know when a detected event occurred, and a measurement to output the bit value with the chosen basis. If the QND measurement results in a detected instance, Bob performs the measurement using the Z or X basis. We are able to delay Bob's measurement choice because the inconclusive outcomes are assumed to be independent of the basis, as explained in section 'Assumptions on Bob's device'. The key point in the virtual protocol is as follows: the security of the events when Alice sends the actual Z basis states and Bob obtains a detected event in the actual protocol with the Z basis can be analysed by imagining that Alice and Bob both employ the X basis to measure respectively the systems A_2 and BE. This means that when Alice sends a virtual state ($c=1$, 2), Bob's measurement basis is always the X basis.

It is clear that the virtual protocol described here is equivalent to the actual protocol in the Supplementary Material (Section I). This is so, because the quantum states sent by Alice are the same in both protocols as well as the announcements made by the two parties. For instance, when Alice sends the virtual states they both measure in the X basis but they announce the Z basis (Step 5). In the actual protocol, these events are used for key generation, and therefore Alice and Bob also announce the Z basis. This means that the protocols are indistinguishable from Eve's perspective as required. It is noteworthy that the virtual protocol does not produce a key; it is merely used for the estimation of the phase error rate.

Azuma's inequality and its application to the security proof

In coherent attacks, Eve interacts with all the signals sent by Alice followed by a joint measurement after listening to all the classical information exchanged between Alice and Bob. In this scenario we use Azuma's inequality[33], which takes into account this dependency and allows us to

derive a relation between the expected values and the observed values. Most importantly, once we have the conditional probabilities on all previous measurement outcomes, we can find the actual number of events observed.

Azuma's inequality can be applied to a stochastic model as long as a sequence of random variables is a martingale and satisfies the bounded difference conditions (BDCs). A Martingale is a sequence of random variables $X^{(0)}$, $X^{(1)}$, …, $X^{(l)}$ for which the expectation $E[\cdot]$ of the next value is equal to the present value in the sequence given that we know all the previous outcomes, i.e., $E[X^{(l+1)}|X^{(0)}, X^{(1)}, \ldots, X^{(l)}] = X^{(l)}$ for all $l \geq 0$. This sequence is said to satisfy BDC if there exists $c^{(l)} > 0$ such that $|X^{(l+1)} - X^{(l)}| \leq c^{(l)}$ for all $l \geq 0$. For N trials of a variable $X^{(l)}$ with $c^{(l)} = 1$, Azuma's inequality states that

$$P[|X^{(N)} - X^{(0)}| > N\delta_{\mathrm{A}}] \leq 2e^{\frac{-N\delta_{\mathrm{A}}^2}{2}}, \tag{40}$$

holds for any $\delta_{\mathrm{A}} \in (0, 1)$. Now, for the lth trial, we define $X^{(l)}$ as

$$X^{(l)} := \Lambda^{(l)} - \sum_{k=1}^{l} P(\zeta_k = 1|\zeta_0, \ldots, \zeta_{k-1}), \tag{41}$$

where $\Lambda^{(l)}$ is a random variable representing the actual number of events (that is $\Lambda^{(l)} = \sum_{k=1}^{l} \zeta_k$) observed during the first l trials, ζ_k is the random variable of interest and it has the value of 0 or 1. Moreover, $P(\zeta_k = 1|\zeta_0, \ldots, \zeta_{k-1})$ is the conditional probability of obtaining the outcome specified by $\zeta_k = 1$ in the kth trial given that the first $k-1$ outcomes are $\zeta_0, \ldots, \zeta_{k-1}$. It is possible to show that the sequence of random variables in Eq. (41) is Martingale and satisfies the BDC. Hence, we can apply the Azuma's inequality and write

$$P[|\Lambda^{(N)} - \sum_{k=1}^{N} P(\zeta_k = 1|\zeta_0, \ldots, \zeta_{k-1})| > N\delta_{\mathrm{A}}] \leq 2e^{\frac{-N\delta_{\mathrm{A}}^2}{2}}, \tag{42}$$

where we use the definition $X^{(0)} = 0$. This also means that

$$\sum_{k=1}^{N} P(\zeta_k = 1|\zeta_0, \ldots, \zeta_{k-1}) - N\delta_{\mathrm{A}} \leq \Lambda^{(N)} \leq \sum_{k=1}^{N} P(\zeta_k = 1|\zeta_0, \ldots, \zeta_{k-1}) + N\delta_{\mathrm{A}}, \tag{43}$$

holds at least with probability $P = 1 - 2e^{\frac{-N\delta_{\mathrm{A}}^2}{2}}$. Therefore,

$$\Lambda^{(N)} = \sum_{k=1}^{N} P(\zeta_k = 1|\zeta_0, \ldots, \zeta_{k-1}) + \delta_{\mathrm{B}}, \tag{44}$$

except for error probability $\epsilon + \hat{\epsilon}$, where the deviation parameter $\delta_{\mathrm{B}} \in [-\Delta, \hat{\Delta}]$. These bounds are defined as $\Delta = f(N, \epsilon)$ and $\hat{\Delta} = f(N, \hat{\epsilon})$ where $f(x, y) = \sqrt{2x \ln(1/y)}$.

Let us now show how we use this inequality in our security proof. In particular, we consider

$$X_{csX}^{(l)} = \Lambda_{csX}^{(l)} - \sum_{k=1}^{l} P(\zeta_{k,csX} = 1|\zeta_0, \ldots, \zeta_{k-1}), \tag{45}$$

where $csX = c, s$ for $c = 1, 2, 3, 4$, as Bob's basis choice is already included in these cases, and $csX = c, s, X$ for $c = 5$. In Eq. (45), $P(\zeta_{k,csX} = 1|\zeta_0, \ldots, \zeta_{k-1})$ is the probability of Alice selecting the state c and Bob observing s (s, X) for $s \in \{0, 1\}$ when $c = 1, 2, 3, 4$ ($c = 5$) in the kth trial, conditional on all the previous outcomes from the measurements $\zeta_0, \ldots, \zeta_{k-1}$. To obtain this probability we first define

$$|\tau\rangle_{\mathrm{SA_1BE}} = |\varphi_{k-1}\rangle_{\mathrm{SA_1BE}}|\varphi_k\rangle_{\mathrm{SA_1BE}}|\varphi_r\rangle_{\mathrm{SA_1BE}}, \tag{46}$$

to be the state prepared by Alice in an execution of the protocol, where $|\varphi_{k-1}\rangle_{\mathrm{SA_1BE}}$, $|\varphi_k\rangle_{\mathrm{SA_1BE}}$ and $|\varphi_r\rangle_{\mathrm{SA_1BE}}$ correspond to all the systems before the kth trial, in the kth trial and in the rest of the trials after k (i.e., $r = N - k$), respectively.

Eve's action can be described as $\hat{U}_{\mathrm{BEE'}}|\tau\rangle_{\mathrm{SA_1BE}}|0\rangle_{E'} = \sum_t \hat{B}_{t\mathrm{B}}|\tau\rangle_{\mathrm{SA_1BE}}|t\rangle_{E'}$, where $\hat{U}_{\mathrm{BEE'}}$ is a unitary transformation acting on systems BEE, $\hat{B}_{t\mathrm{B}}$ is the Kraus operator, which acts on system BE depending on Eve's measurement outcome t, and $|t\rangle_{\{t=1,\ 2,\ \ldots\}}$ is an orthonormal basis. It is noteworthy that here we use the subscript E to refer to Eve's system originating from a THA and E' corresponds to an additional ancilla system in her hands. Alice and Bob only communicate after performing the measurements so these parameters are independent of the state preparation.

In order to consider Alice's and Bob's measurements previous to the kth trial, we define the operator $\hat{\mathcal{O}}_{k-1,\mathrm{SBE}} = \otimes_{\nu=1}^{k-1} \hat{M}_{S_\nu \mathrm{BE}_\nu}$, where $\hat{M}_{S_\nu \mathrm{BE}_\nu}$ denotes the Kraus operator associated with the νth measurement outcome of Alice and Bob. Hence, after Eve's interaction, the normalised kth state of the system SBE conditioned on the measurement outcomes, O_{k-1}, and the detected event can be expressed as

$$\hat{\rho}_{k|O_{k-1}}^{\mathrm{SBE}} = \frac{\hat{\sigma}_{k|O_{k-1}}^{\mathrm{SBE}}}{\mathrm{Tr}\left(\hat{\sigma}_{k|O_{k-1}}^{\mathrm{SBE}}\right)}, \tag{47}$$

where the state $\hat{\sigma}_{k|O_{k-1}}^{\mathrm{SBE}}$ is defined shortly below (see Eq. (49)). We know that

$$\hat{\sigma}_{k|O_{k-1}}^{\mathrm{SA_1BE}} = \sum_t \mathrm{Tr}_{\bar{k}}\left[\hat{F}_{\mathrm{BE}_k}\hat{\mathcal{O}}_{k-1,\mathrm{SBE}}\hat{B}_{t\mathrm{B}}|\tau\rangle_{\mathrm{SA_1BE}}\langle\tau|\hat{B}_{t\mathrm{B}}^{\dagger}\hat{\mathcal{O}}_{k-1,\mathrm{SBE}}^{\dagger}\hat{F}_{\mathrm{BE}_k}^{\dagger}\right], \tag{48}$$

where $\mathrm{Tr}_{\bar{k}}$ is the partial trace over the systems S, A_1 and BE for all the events that are not in the kth trial, and $\hat{F}_{\mathrm{BE}_k}$ is Bob's Kraus operator acting on the kth system, corresponding to the detected events. This means taking the trace with the basis $\{|\vec{x}_{k-1}\rangle, |\vec{x}_r\rangle\}$, where $|\vec{x}_{k-1}\rangle$ corresponds to all the systems in the first $k-1$ runs and $|\vec{x}_r\rangle$ to the rest of the systems after k. Then, we can rewrite Eq. (48) as

$$\hat{\sigma}_{k|O_{k-1}}^{\mathrm{SBE}} = \sum_t \sum_{\vec{x}_{k-1},\vec{x}_r} \mathrm{Tr}_{\mathrm{A_1}}^k\left[A_{t,\mathrm{BE}|O_{k-1}}^{(\vec{x}_{k-1},\vec{x}_r)}|\varphi_k\rangle_{\mathrm{SA_1BE}}\langle\varphi_k|A_{t,\mathrm{BE}|O_{k-1}}^{\dagger(\vec{x}_{k-1},\vec{x}_r)}\right], \tag{49}$$

where $\mathrm{Tr}_{\mathrm{A_1}}^k$ is the partial trace over the system A_1 in the kth trial and $A_{t,\mathrm{BE}|O_{k-1}}^{(\vec{x}_{k-1},\vec{x}_r)}$ is the Kraus operator acting on the kth system conditional on all the previous detected events, and it is defined as

$$A_{t,\mathrm{BE}|O_{k-1}}^{(\vec{x}_{k-1},\vec{x}_r)} = \langle\vec{x}_r|\langle\vec{x}_{k-1}|\hat{F}_{\mathrm{BE}_k}\hat{O}_{k-1,\mathrm{SBE}}\hat{B}_{t\mathrm{B}}|\varphi_{k-1}\rangle_{\mathrm{SA_1BE}}|\varphi_r\rangle_{\mathrm{SA_1BE}}. \tag{50}$$

By substituting Eq. (37) into Eq. (49) we get

$$\hat{\sigma}_{k|O_{k-1}}^{\mathrm{SBE}} = \sum_{c,c\prime}\sqrt{P(c)P(c\prime)}\sum_t\sum_{\vec{x}_{k-1},\vec{x}_r}\mathrm{Tr}_{\mathrm{A_1}}^k\left[A_{t,\mathrm{BE}|O_{k-1}}^{(\vec{x}_{k-1},\vec{x}_r)}|c\rangle_S\langle c\prime| \otimes |\vartheta^{(c)}\rangle_{\mathrm{A_1BE}}\langle\vartheta^{(c\prime)}|A_{t,\mathrm{BE}|O_{k-1}}^{\dagger(\vec{x}_{k-1},\vec{x}_r)}\right]. \tag{51}$$

It is clear now that this state is dependent on Eve's action as well as on the previous outcomes. Also, note that the partial trace only acts on system A_1. The probability that Alice obtains the outcome c, Bob selects the X basis and obtains a bit value s conditional on all the previous measurement outcomes is calculated as

$$\begin{aligned} P_{csX|O_{k-1}} &= \frac{P(X\cap c)}{\mathrm{Tr}\left(\hat{\sigma}_{k|O_{k-1}}^{\mathrm{SBE}}\right)}\sum_t\sum_{\vec{x}_{k-1},\vec{x}_r}\mathrm{Tr}\left[A_{t,\mathrm{BE}|O_{k-1}}^{(\vec{x}_{k-1},\vec{x}_r)}\mathrm{Tr}_{A_1}^k\left[|\vartheta^{(c)}\rangle\langle\vartheta^{(c)}|_{\mathrm{A_1BE}}\right]A_{t,\mathrm{BE}|O_{k-1}}^{\dagger(\vec{x}_{k-1},\vec{x}_r)}\hat{M}_{sX}\right] \\ &= \frac{P(X\cap c)}{\mathrm{Tr}(\hat{\sigma}_{k|O_{k-1}}^{\mathrm{SBE}})}\mathrm{Tr}\left[\hat{D}_{sX|O_{k-1}}\mathrm{Tr}_{A_1T}^k\left[|\vartheta^{(c)}\rangle\langle\vartheta^{(c)}|_{\mathrm{A_1BE}}\right]\right] \end{aligned} \tag{52}$$

where $P(X \cap c) = P(c)$ for $c = 1, 2, 3, 4$ and $P(X \cap c) = P(c)P(X_B)$ for $c = 5$. In this expression, $\hat{D}_{sX|O_{k-1}} = \sum_t\sum_{\vec{x}_{k-1},\vec{x}_r} A_{t,\mathrm{BE}|O_{k-1}}^{\dagger(\vec{x}_{k-1},\vec{x}_r)}\hat{M}_{sX}A_{t,\mathrm{BE}|O_{k-1}}^{(\vec{x}_{k-1},\vec{x}_r)}$ represents Eve's action as well as Bob's measurement. This is independent of c, which means that Eve cannot behave differently depending of the state sent. Importantly, the probability $P_{csX|O_{k-1}}$ essentially corresponds to the actual yields $Y_{sX,j\beta}$ in section 'Estimation of the phase error rate' when $c = 3, 4, 5$. It is noteworthy that the yields in this section are normalised by the detected events while those in the Results section are not. In the finite key size regime, the normalisation according to the detected events results in a better performance; however, in the limit of large number of pulses, they are essentially the same. As we consider this limit throughout this paper, in section 'Estimation of the phase error rate', we adopt the yields that are not normalised by the detected events for simplicity of explanation. We know Λ_{3sX}, Λ_{4sX} and Λ_{5sX} by collecting the corresponding number of events from the actual protocol. Therefore, using Azuma's inequality, i.e., Eq. (44), we can calculate the conditional probabilities that correspond to the yields $Y_{sX,0Z}^{(Z)}$, $Y_{sX,1Z}^{(Z)}$ and $Y_{sX,0X}^{(X)}$, respectively. From section 'Estimation of the phase error rate', we know how these yields are related to the transmission rates and, in turn, how these are related to the virtual yields $Y_{1X,0Z}^{(Z)vir}$ and $Y_{0X,1Z}^{(Z)vir}$. Here we would like to emphasise that using Eq. (18) we can calculate these yields, which correspond to the probabilities $P_{11X|O_{k-1}}$ and $P_{20X|O_{k-1}}$, respectively, both of which are conditional on the previous measurement outcomes. Using Azuma's inequality again, we can find the number of number of events, Λ_{11X} and Λ_{20X}, which are the number of phase errors, and this concludes the estimation of the phase error rate.

Coefficients

In this section, we list the coefficients used in section 'Estimation of the phase error rate'. Direct calculations show that the coefficients A_j, B_j and C_j

for Eqs. (14)–(18) are given by

$$A_j = \frac{1}{4}\left[|a_{0Z}|^2 + (-1)^j\left(a^*_{0Z}a_{1Z}\langle\phi_{0Z}|\phi_{1Z}\rangle_{BE} + a_{0Z}a^*_{1Z}\langle\phi_{1Z}|\phi_{0Z}\rangle_{BE}\right) + |a_{1Z}|^2\right],$$
$$B_j = \frac{1}{4}\sqrt{|a_{0Z}|^2 + (-1)^j\left(a^*_{0Z}a_{1Z}\langle\phi_{0Z}|\phi_{1Z}\rangle_{BE} + a_{0Z}a^*_{1Z}\langle\phi_{1Z}|\phi_{0Z}\rangle_{BE}\right) + |a_{1Z}|^2}$$
$$\times\sqrt{|b_{0Z}|^2 + (-1)^j\left(b^*_{0Z}b_{1Z}\langle\phi^\perp_{0Z}|\phi^\perp_{1Z}\rangle_{BE} + b_{0Z}b^2_{1Z}\langle\phi^\perp_{1Z}|\phi^\perp_{0Z}\rangle_{BE}\right) + |b_{1Z}|^2},$$
$$C_j = \frac{1}{4}\left[|b_{0Z}|^2 + (-1)^j\left(b^*_{0Z}b_{1Z}\langle\phi^\perp_{0Z}|\phi^\perp_{1Z}\rangle_{BE} + b_{0Z}b^*_{1Z}\langle\phi^\perp_{1Z}|\phi^\perp_{0Z}\rangle_{BE}\right) + b^2_{1Z}\right]. \quad (53)$$

DATA AVAILABILITY

No datasets were generated or analysed during the current study.

ACKNOWLEDGEMENTS

We thank Koji Azuma, Hoi-Kwong Lo, Go Kato, Norbert Lütkenhaus, Masato Koashi, Toshihiko Sasaki, Akihiro Mizutani, Guillermo Currás Lorenzo and Weilong Wang for very valuable discussions. This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO), the Fondo Europeo de Desarrollo Regional (FEDER) through grants TEC2014-54898-R and TEC2017-88243-R, and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 675662 (project QCALL). K.T. acknowledges support from JSPS KAKENHI Grant Numbers JP18H05237 18H05237, ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan), and JST-CREST JPMJCR 1671.

AUTHOR CONTRIBUTIONS

All authors discussed the main idea and M.P. performed the analytical calculations and the numerical simulations. All authors analysed the results and prepared the manuscript.

ADDITIONAL INFORMATION

Supplementary Information accompanies the paper on the *npj Quantum Information* website (https://doi.org/10.1038/s41534-019-0180-9).

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Gisin, N., Ribordy, G., Tittel, W. & Zbinden, H. Quantum cryptography. *Rev. Mod. Phys.* **74**, 145 (2002).
2. Scarani, V. et al. The security of practical quantum key distribution. *Rev. Mod. Phys.* **81**, 1301 (2009).
3. Lo, H.-K., Curty, M. & Tamaki, K. Secure quantum key distribution. *Nat. Photonics* **8**, 595–604 (2014).
4. Lydersen, L. et al. Hacking commercial quantum cryptographic systems by tailored bright illumination. *Nat. Photonics* **4**, 686–698 (2010).
5. Gerhardt, I. et al. Full-field implementation of a perfect eavesdropper on a quantum cryptography system. *Nat. Commun.* **2**, 349 EP – (2011).
6. Lo, H.-K., Curty, M. & Qi, B. Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.* **108**, 130503 (2012).
7. Rubenok, A., Slater, J. A., Chan, P., Lucio-Martinez, I. & Tittel, W. Real-world two-photon interference and proof-of-principle quantum key distribution immune to detector attacks. *Phys. Rev. Lett.* **111**, 130501 (2013).
8. da Silva, T. F. et al. Proof-of-principle demonstration of measurement-device-independent quantum key distribution using polarization qubits. *Phys. Rev. A* **88**, 052303 (2013).
9. Liu, Y. et al. Experimental measurement-device-independent quantum key distribution. *Phys. Rev. Lett.* **111**, 130502 (2013).
10. Tang, Z. et al. Experimental demonstration of polarization encoding measurement-device-independent quantum key distribution. *Phys. Rev. Lett.* **112**, 190503 (2014).
11. Yin, H.-L. et al. Measurement-device-independent quantum key distribution over a 404 km optical fiber. *Phys. Rev. Lett.* **117**, 190501 (2016).
12. Roberts, G. L. et al. Experimental measurement-device-independent quantum digital signatures. *Nat. Commun.* **8**, 1098 (2017).
13. Hwang, W.-Y. Quantum key distribution with high loss: toward global secure communication. *Phys. Rev. Lett.* **91**, 057901 (2003).
14. Lo, H.-K., Ma, X. & Chen, K. Decoy state quantum key distribution. *Phys. Rev. Lett.* **94**, 230504 (2005).
15. Wang, X.-B. Beating the photon-number-splitting attack in practical quantum cryptography. *Phys. Rev. Lett.* **94**, 230503 (2005).
16. Gottesman, D., Lo, H.-K., Lütkenhaus, N. & Preskill, J. Security of quantum key distribution with imperfect devices. *Quantum Inf. Comput.* **7**, 431 (2007).
17. Tamaki, K., Curty, M., Kato, G., Lo, H.-K. & Azuma, K. Loss-tolerant quantum cryptography with imperfect sources. *Phys. Rev. A* **90**, 052314 (2014).
18. Boaron, A. et al. Secure quantum key distribution over 421 km of optical fiber. *Phys. Rev. Lett.* **121**, 190502 (2018).
19. Gisin, N., Fasel, S., Kraus, B., Zbinden, H. & Ribordy, G. Trojan-horse attacks on quantum-key-distribution systems. *Phys. Rev. A* **73**, 022320 (2006).
20. Vakhitov, A., Makarov, V. & Hjelme, D. R. Large pulse attack as a method of conventional optical eavesdroppong in quantum cryptography. *J. Mod. Opt.* **48**, 2023 (2001).
21. Lucamarini, M., Choi, I., Ward, M. B., Yuan, J. F. D. Z. L. & Shields, A. J. Practical security bounds against the trojan-horse attack in quantum key distribution. *Phys. Rev. X* **5**, 031030 (2015).
22. Tamaki, K., Curty, M. & Lucamarini, M. Decoy-state quantum key distribution with a leaky source. *New J. Phys.* **18**, 065008 (2016).
23. Wang, W., Tamaki, K. & Curty, M. Finite-key security analysis for quantum key distribution with leaky sources. *New J. Phys.* **20**, 083027 (2018).
24. Mizutani, A. et al. Quantum key distribution with setting-choice-independently correlated light sources. *npj Quantum Inf.* **5**, 8 (2019).
25. Bruss, D. Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.* **81**, 3018–3021 (1998).
26. Inoue, K., Waks, E. & Yamamoto, Y. Differential phase shift quantum key distribution. *Phys. Rev. Lett.* **89**, 037902 (2002).
27. Takesue, H. et al. Quantum key distribution over a 40-db channel loss using superconducting single-photon detectors. *Nat. Photonics* **1**, 343 (2007).
28. Stucki, D. et al. High rate, long-distance quantum key distribution over 250km of ultra low loss fibres. *New J. Phys.* **11**, 075003 (2009).
29. Lo, H.-K. & Preskill, J. Security of quantum key distribution using weak coherent states with nonrandom phases. *Quantum Inf. Comput.* **7**, 431–458 (2007).
30. Nagamatsu, Y. et al. Security of quantum key distribution with light sources that are not independently and identically distributed. *Phys. Rev. A* **93**, 042325 (2016).
31. Koashi, M. Complementarity, distillable secret key, and distillable entanglement. *arXiv:0704.3661* (2007).
32. Koashi, M. Simple security proof of quantum key distribution based on complementarity. *New J. Phys.* **11**, 045018 (2009).
33. Azuma, K. Weighted sums of certain dependent random variables. *Tohoku Math. J.* **19**, 357–367 (1967).
34. Xu, F. et al. Experimental quantum key distribution with source flaws. *Phys. Rev. A* **92**, 032305 (2015).
35. Xu, F., Qi, B. & Lo, H.-K. Experimental demonstration of phase-remapping attack in a practical quantum key distribution system. *New J. Phys.* **12**, 113026 (2010).
36. Jiang, M.-S., Sun, S.-H., Li, C.-Y. & Liang, L.-M. Frequency shift attack on 'plug-and-play' quantum key distribution systems. *J. Mod. Opt.* **61**, 147–153 (2014).
37. Mynbaev, D. K. & Scheiner, L. L. *Fiber-optic communications technology* (Prentice Hall, 2001).
38. Bennett, C.-H. & Brassard, G. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, (vol. 175), 175–179 (1984).
39. Honjo, T., Inoue, K. & Takahashi, H. Differential-phase-shift quantum key distribution experiment with a planar light-wave circuit mach–zehnder interferometer. *Opt. Lett.* **29**, 2797–2799 (2004).
40. Li, G. Recent advances in coherent optical communication. *Adv. Opt. Photon.* **1**, 279–307 (2009).
41. Tang, Z., Wei, K., Bedroya, O., Qian, L. & Lo, H.-K. Experimental measurement-device-independent quantum key distribution with imperfect sources. *Phys. Rev. A* **93**, 042308 (2016).
42. Wang, Y., Primaatmaja, I. W., Lavie, E., Varvitsiotis, A. & Lim, C. C. W. Characterising the correlations of prepare-and-measure quantum networks. *npj Quantum Inf.* **5**, 17 (2019).
43. Coles, P. J., Metodiev, E. M. & Lütkenhaus, N. Numerical approach for unstructured quantum key distribution. *Nat. Commun.* **7**, 11712 (2016).

44. Winick, A., Lütkenhaus, N. & Coles, P. J. Reliable numerical key rates for quantum key distribution. *Quantum* **2**, 77 (2018).
45. Shor, P. W. & Preskill, J. Simple proof of security of the bb84 quantum key distribution protocol. *Phys. Rev. Lett.* **85**, 441–444 (2000).
46. Mayers, D. In *Advances in Cryptology — CRYPTO '96* (ed. Koblitz, N.) 343–357 (Springer, Berlin Heidelberg, 1996).