Security of quantum key distribution with intensity correlations

Víctor Zapatero[1], Álvaro Navarrete[1], Kiyoshi Tamaki[2], and Marcos Curty[1]

[1]Escuela de Ingeniería de 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

The decoy-state method in quantum key distribution (QKD) is a popular technique to approximately achieve the performance of ideal single-photon sources by means of simpler and practical laser sources. In high-speed decoy-state QKD systems, however, intensity correlations between succeeding pulses leak information about the users' intensity settings, thus invalidating a key assumption of this approach. Here, we solve this pressing problem by developing a general technique to incorporate arbitrary intensity correlations to the security analysis of decoy-state QKD. This technique only requires to experimentally quantify two main parameters: the correlation range and the maximum relative deviation between the selected and the actually emitted intensities. As a side contribution, we provide a non-standard derivation of the asymptotic secret key rate formula from the non-asymptotic one, in so revealing a necessary condition for the significance of the former.

1 Introduction

Quantum key distribution [1, 2, 3] (QKD) is a technique that enables secure and remote delivery of cryptographic keys based on the laws of quantum mechanics. The interest of QKD is that, when combined with the one-time-pad encryption scheme [4], it allows for information-theoretically secure communication, unconcerned about the capabilities of future adversaries and the progress of classical or even quantum computers. For this reason, since its conception in 1984 [5], QKD has experienced a tremendous development both in theory and in practice, in so becoming a commercial technology that represents the most mature application of quantum information science. Nevertheless, various challenges must still be addressed in order to achieve the widespread adoption of QKD.

In real-life implementations, the information carrier of QKD is the quantum of light or photon, and due to the low transmissivity of single photons in typical optical channels —which, for instance, in the case of optical fibers decreases exponentially with the fiber length [6, 7, 8, 9]— one major challenge consists of achieving high secret key generation rates at long distances. For this purpose, one natural approach is to increase the repetition rate of the laser source in the transmitter station. However, for repetition rates of the order of GHz, it has been shown that intensity correlations between succeeding pulses appear [10, 11], potentially opening a security loophole.

Víctor Zapatero: vzapatero@com.uvigo.es

Álvaro Navarrete: anavarrete@com.uvigo.es

To be precise, in the absence of ideal single-photon sources, most QKD protocols use simpler laser sources that operate emitting phase-randomised weak coherent pulses (PRWCPs). This is so because PRWCPs allow the QKD users to implement the so-called decoy-state method [12, 13, 14, 15], a technique to tightly lower-bound the extractable secret key length of a QKD session. Importantly, standard decoy-state analyses rely on a fundamental assumption: for any given signal, its detection probability (or so-called "yield") conditioned on the emission of a certain number of photons does not depend on the intensity of the pulse, *i.e.,* on its mean photon number. Nevertheless, this assumption fails in the presence of a side channel leaking information about the intensity setting [16], and for GHz (or higher frequency) repetition rates one such side channel is represented by intensity correlations. Intuitively, the eavesdropper (Eve) could exploit the correlations to gain information about previous intensity settings, which would allow her to make the n-photon yields dependent on them.

This being the case, and aiming to develop ultrafast decoy-state-based QKD systems, the question arises of how to account for arbitrary intensity correlations in the security analysis. In this regard, the existing security proofs are notably restricted. For instance, preliminary results presented in [17, 18] deal with setting-choice-independent correlations, which neglect the possibility of information leakage and hence do not cover the major threat. On the other hand, the authors of [10] go beyond setting-choice-independent correlations by providing a post-processing hardware countermeasure whose application is limited to particular instances of nearest neighbors intensity correlations. Similarly, it is worth mentioning the progress reported in [19] to develop an intensity modulator (IM) that mitigates the effect of intensity correlations. Lastly, the recent work in [20] presents a general technique to accomodate various other device imperfections, but it does not incorporate the use of the decoy-state method.

In this work, we provide the missing security analysis for decoy-state QKD with arbitrary intensity correlations. Despite asymptotic, our approach is experimental-friendly in the following sense. In the first place, it only requires to upper bound two parameters presumably easy to quantify in an experiment (see for instance [11]): the correlation range and the maximum relative deviation between the selected intensity settings and the actually emitted intensities (which do not match in general due to the correlations). In the second place, it allows for improved secret key rates in case a specific correlations model is known to describe the IM.

As a side contribution, we use elementary results in statistical convergence to rigorously justify the asymptotic secret key rate formula from the non-asymptotic one. Crucially, the justification relies on a very natural necessary condition on the observables. Whenever not guaranteed by some special symmetry of the protocol (such as the delayed setting choice that enables a counterfactual scenario in the absence of information leakage), this condition must be taken as an assumption, in which case it restricts the capabilities of the adversary. In particular, the formula tolerates a restricted type of coherent attacks that we characterize in detail.

The structure of the paper goes as follows. There are six Results subsections. In the first two, Sec. 2.1 and Sec. 2.2, we present the general physical assumptions we impose on the intensity correlations and provide a method to quantify their effect in the parameter estimation procedure. This procedure is explained in full detail in Sec. 2.3 and Sec. 2.4. In Sec. 2.5 we establish the asymptotic key rate formula and discuss the necessary condition on which it relies. The last Results subsection, Sec. 2.6, is devoted to evaluate the rate-distance performance of our method for different values of the two parameters that characterize the correlations. In the Discussion section, Sec. 3, we summarize the con-

tributions and limitations of our work, commenting on possible future directions. Lastly, Sec. 4.1, Sec. 4.2 and Sec. 4.3 are the Methods subsections, which include all the necessary technical derivations that support our results. The appendices attached at the end include a description of the channel model we use for the simulations, together with some complementary results.

2 Results

For illustration purposes, we consider a standard polarization encoding decoy-state BB84 protocol [15], although our results can be readily extended to any QKD protocol that relies on decoy-states. In each round k, with $k = 1, \ldots, N$, the sender (Alice) selects a basis $x_k \in M = \{\mathrm{X}, \mathrm{Z}\}$ with probability q_{x_k}, a uniform raw key bit $r_k \in \mathbb{Z}_2 = \{0, 1\}$, and an intensity setting $a_k \in A = \{\mu, \nu, \omega\}$ with probability p_{a_k} and $\mu > \nu > \omega \geq 0$. Note that the values of the probabilities q_{x_k} and p_{a_k} respectively depend on the basis and intensity settings only, but not on the particular round k. Then, she encodes the BB84 state defined by x_k and r_k in a PRWCP with intensity setting a_k, and sends it to the receiver (Bob) through the quantum channel. Importantly, as explained below, the actual mean photon number of the pulse might not match the setting a_k due to the intensity correlations. Furthermore, regarding the transmitter, we assume perfect phase randomization, no state preparation flaws and no side-channels for simplicity. Also, possible intensity correlations in the qubit encoding (which may arise, *e.g.*, when using time-bin encoding) are neglected in this work.

Similarly, Bob selects a basis $y_k \in M$ with probability q_{y_k} (whose value, again, does not depend on the round k) and performs a positive operator-valued measure (POVM) on the incident pulse, given by $\{\hat{M}_{B_k}^{y_k,s_k}\}_{s_k \in \{0,1,f\}}$. Here, B_k denotes Bob's k-th incoming pulse, s_k stands for Bob's classical outcome and f stands for "no click". As usual, the basis-independent detection efficiency condition is assumed, such that $\hat{M}_{B_k}^{\mathrm{Z},f} = \hat{M}_{B_k}^{\mathrm{X},f}$. Thus, we shall simply denote these two operators by $\hat{M}_{B_k}^{f}$. Note that this assumption could be removed by the use of measurement-device-independent (MDI) QKD [21] or twin-field QKD [22].

2.1 Characterizing the intensity correlations

Let us denote the record of intensity settings up to round k by $\vec{a}_k = a_1, a_2, \ldots, a_k$ —where $a_j \in A$ for all j— and let α_k denote the actual intensity delivered in round k. In full generality, α_k is a continuous random variable whose probability density function is fixed by the record of settings $\vec{a}_k$. This function, which we denote as $g_{\vec{a}_k}(\alpha_k)$, is referred to as the correlations model. Below, we list three assumptions about the intensity correlations on which our analysis relies.

Assumption 1. As supported by GHz-clock QKD experiments [10, 11], we shall consider that the correlations do not compromise the poissonian character of the photon number statistics of the source conditioned on the value of the actual intensity, α_k. That is to say, for any given round k, and for all $n_k \in \mathbb{N}$,

$$p(n_k|\alpha_k) = \frac{e^{-\alpha_k}\alpha_k^{n_k}}{n_k!}. \tag{1}$$

Assumption 2. For all possible records $\vec{a}_k$, we shall assume that

$$\left|1-\frac{\alpha_k}{a_k}\right| \leq \delta_{\max}. \tag{2}$$

Namely, for every round, we impose that $g_{\vec{a}_k}(\alpha_k)$ is only nonzero for $\alpha_k \in [a_k^-, a_k^+]$, where $a_k^\pm = a_k(1\pm\delta_{\max})$. Thus, $\delta_{\max}$ defines the maximum relative deviation between a_k and α_k. Note that we are assuming here that the value of $\delta_{\max}$ does not depend on a_k for simplicity, but such dependence could be easily incorporated in the analysis to obtain slightly tighter results. Also, we remark that a bound of the type of Eq. (2) has been quantified in a recent experiment reported in [11].

From Eq. (1) and Eq. (2), it follows that the photon number statistics of round k read

$$p_{n_k|\vec{a}_k} = \int_{a_k^-}^{a_k^+} g_{\vec{a}_k}(\alpha_k)\frac{e^{-\alpha_k}\alpha_k^{n_k}}{n_k!}\,d\alpha_k, \tag{3}$$

for all $n_k \in \mathbb{N}$.

Assumption 3. We assume that the intensity correlations have a finite range, say ξ, meaning that $g_{\vec{a}_k}(\alpha_k)$ is independent of those previous settings a_j with $k-j>\xi$.

Beyond the assumptions presented here, we shall consider that $g_{\vec{a}_k}(\alpha_k)$ is unknown, such that our results are model-independent.

2.2 Quantifying the effect of the intensity correlations

Here, we rely on the three assumptions introduced in Sec. 2.1 to account for the effect of intensity correlations in the decoy-state analysis. A key idea —originally presented in [16] to deal with Trojan horse attacks— is to pose a restriction on the maximum bias that Eve can induce between the n-photon yields associated to different intensity settings. For this purpose, we use a fundamental result presented in [23] and further developed in [20]. Since this result is a direct consequence of the Cauchy–Schwarz (CS) inequality in complex Hilbert spaces, we shall refer to it as the CS constraint. The reader is referred to the Methods Sec. 4.1 for a definition of this result, and below we present the relevant restrictions we derive with it.

Precisely, for any given round k, photon number $n \in \mathbb{N}$, intensity setting $c \in A$ and bit value $r \in \{0,1\}$, we define the yield $Y_{n,c}^{(k)} = p^{(k)}(\text{click}|n,c,\text{Z},\text{Z})$ and the error probability $H_{n,c,r}^{(k)} = p^{(k)}(\text{err}|n,c,\text{X},\text{X},r)$, where the right-hand sides are shorthand for $p(s_k \neq f|n_k = n, a_k = c, x_k = \text{Z}, y_k = \text{Z})$ and $p(s_k \neq f, s_k \neq r_k|n_k = n, a_k = c, x_k = \text{X}, y_k = \text{X}, r_k = r)$, respectively. Then, a major result of this work is to show that, for any two distinct intensity settings a and b, their yields and error probabilities satisfy

$$\begin{aligned} G_-\left(Y_{n,a}^{(k)}, \tau_{ab,n}^{\xi}\right) \leq Y_{n,b}^{(k)} \leq G_+\left(Y_{n,a}^{(k)}, \tau_{ab,n}^{\xi}\right), \\ G_-\left(H_{n,a,r}^{(k)}, \tau_{ab,n}^{\xi}\right) \leq H_{n,b,r}^{(k)} \leq G_+\left(H_{n,a,r}^{(k)}, \tau_{ab,n}^{\xi}\right), \end{aligned} \tag{4}$$

for all k, n and r, where

$$G_-(y,z) = \begin{cases} g_-(y,z) & \text{if} \quad y > 1-z \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad G_+(y,z) = \begin{cases} g_+(y,z) & \text{if} \quad y < z \\ 1 & \text{otherwise} \end{cases} \tag{5}$$

with $g_\pm(y,z) = y+(1-z)(1-2y) \pm 2\sqrt{z(1-z)y(1-y)}$, ξ stands for the correlation range and

$$\tau^{\xi}_{ab,n} = \begin{cases} e^{a^- + b^- - (a^+ + b^+)} \left[1 - \sum_{c \in A} p_c \left(e^{-c^-} - e^{-c^+}\right)\right]^{2\xi} & \text{if} \quad n = 0, \\ e^{a^+ + b^+ - (a^- + b^-)} \left(\frac{a^- b^-}{a^+ b^+}\right)^n \left[1 - \sum_{c \in A} p_c \left(e^{-c^-} - e^{-c^+}\right)\right]^{2\xi} & \text{if} \quad n \geq 1. \end{cases} \tag{6}$$

The full derivation of this result is given in the Methods Sec. 4.1.

Notably, Eq. (4) must be combined with a decoy-state method in order to estimate the numbers of counts and errors triggered by single photon emissions, which determine the secret key rate. For this purpose, in Sec. 2.3 we provide a decoy-state analysis that relies on assumptions 1 and 2 to deal with intensity correlations, and in Sec. 2.4 we present the resulting linear programs that fulfill the parameter estimation. In this regard, since the constraints of Eq. (4) are non-linear, first-order approximations with respect to some reference parameters are derived from them, which we refer to as the linearized CS constraints. Importantly, replacing the original functions by their linear approximations leads to looser but valid constraints too, thanks to the convexity of these functions.

2.3 Decoy-state method

The Z basis gain with intensity setting a is defined as $Z_{a,N} = \sum_{k=1}^{N} Z_a^{(k)}$ with $Z_a^{(k)} = \mathbb{1}_{\{a_k = a, x_k = y_k = \mathrm{Z}, s_k \neq f\}}$. That is to say, $Z_a^{(k)} = 1$ if, in round k, both parties select the Z basis, Alice selects intensity setting a and a click occurs, and zero otherwise. Thus,

$$\langle Z_a^{(k)} \rangle = p^{(k)}(a, \mathrm{Z}, \mathrm{Z}, \text{click}) = q_{\mathrm{Z}}^2 p_a \sum_{n=0}^{\infty} p^{(k)}(n, \text{click}|a, \mathrm{Z}, \mathrm{Z}) = q_{\mathrm{Z}}^2 p_a \sum_{n=0}^{\infty} p^{(k)}(n|a) Y_{n,a}^{(k)}, \tag{7}$$

where the yield $Y_{n,a}^{(k)}$ is defined in Sec. 2.2 and one can generically refer to the $\langle Z_a^{(k)} \rangle$ as the "expected gains of round k". Going back to Eq. (7), note that

$$p^{(k)}(n|a) = \sum_{\vec{a}_{k-1}} p_{a_1} \cdots p_{a_{k-1}} p_n|_{a, \vec{a}_{k-1}}, \tag{8}$$

and in virtue of Eq. (3), the record-independent bounds

$$p^{(k)}(0|a) \in \left[e^{-a^+}, e^{-a^-}\right] \quad \text{and} \quad p^{(k)}(n|a) \in \left[\frac{e^{-a^-} a^{-\,n}}{n!}, \frac{e^{-a^+} a^{+\,n}}{n!}\right] \ (n \geq 1) \tag{9}$$

follow from the decreasing (increasing) character of $e^{-x}x^n$ for $n = 0$ ($n \geq 1$) in the interval $x \in (0,1)$. Explicitly using these intervals in Eq. (7), one obtains

$$\frac{\langle Z_a^{(k)} \rangle}{q_{\mathrm{Z}}^2 p_a} \geq e^{-a^+} Y_{0,a}^{(k)} + \sum_{n=1}^{\infty} \frac{e^{-a^-} a^{-\,n}}{n!} Y_{n,a}^{(k)} \quad \text{and} \quad \frac{\langle Z_a^{(k)} \rangle}{q_{\mathrm{Z}}^2 p_a} \leq e^{-a^-} Y_{0,a}^{(k)} + \sum_{n=1}^{\infty} \frac{e^{-a^+} a^{+\,n}}{n!} Y_{n,a}^{(k)} \tag{10}$$

for all $a \in A$ and $k = 1, \ldots, N$. Further selecting a threshold photon number for the numerics, n_{cut}, from Eq. (10) we have

$$\begin{aligned} &\frac{\langle Z_a^{(k)} \rangle}{q_{\mathrm{Z}}^2 p_a} \geq e^{-a^+} Y_{0,a}^{(k)} + \sum_{n=1}^{n_{\mathrm{cut}}} \frac{e^{-a^-} a^{-\,n}}{n!} Y_{n,a}^{(k)} \quad \text{and} \\ &\frac{\langle Z_a^{(k)} \rangle}{q_{\mathrm{Z}}^2 p_a} \leq 1 - e^{-a^+} + e^{-a^-} Y_{0,a}^{(k)} - \sum_{n=1}^{n_{\mathrm{cut}}} \frac{e^{-a^+} a^{+\,n}}{n!} \left(1 - Y_{n,a}^{(k)}\right) \end{aligned} \tag{11}$$

for all $a \in A$ and $k = 1, \ldots, N$, where in the second inequality we have used the fact that $\sum_{n=n_{\rm cut}+1}^{\infty} Y_{n,a}^{(k)} e^{-a^+} a^{+\,n}/n! \leq 1 - \sum_{n=0}^{n_{\rm cut}} e^{-a^+} a^{+\,n}/n!$. Importantly, replacing Z by X everywhere, one obtains the corresponding analysis for the X basis gains and yields of round k.

On the other hand, similar constraints can be imposed on the error counts. To be precise, the number of X basis error counts with setting a is defined as $E_a = \sum_{k=1}^{N} E_a^{(k)}$ with $E_a^{(k)} = X_a^{(k)} \mathbb{1}_{\{r_k \neq s_k\}}$, such that

$$\langle E_a^{(k)} \rangle = p^{(k)}(a, \mathrm{X}, \mathrm{X}, \mathrm{err}) = q_{\mathrm{X}}^2 p_a \sum_{n=0}^{\infty} p^{(k)}(n, \mathrm{err}|a, \mathrm{X}, \mathrm{X}) = q_{\mathrm{X}}^2 p_a \sum_{n=0}^{\infty} p^{(k)}(n|a) H_{n,a}^{(k)}, \quad (12)$$

where we defined $H_{n,a}^{(k)} = p^{(k)}(\mathrm{err}|n, a, \mathrm{X}, \mathrm{X}) = (H_{n,a,0}^{(k)} + H_{n,a,1}^{(k)})/2$. Now, making use of Eq. (9) as before and selecting a threshold photon number $n_{\rm cut}$, it follows that

$$\begin{aligned} \frac{\langle E_a^{(k)} \rangle}{q_{\mathrm{X}}^2 p_a} &\geq e^{-a^+} H_{0,a}^{(k)} + \sum_{n=1}^{n_{\rm cut}} \frac{e^{-a^-} a^{-\,n}}{n!} H_{n,a}^{(k)} \quad \text{and} \\ \frac{\langle E_a^{(k)} \rangle}{q_{\mathrm{X}}^2 p_a} &\leq 1 - e^{-a^+} + e^{-a^-} H_{0,a}^{(k)} - \sum_{n=1}^{n_{\rm cut}} \frac{e^{-a^+} a^{+\,n}}{n!} \left(1 - H_{n,a}^{(k)}\right) \end{aligned} \quad (13)$$

for all $a \in A$ and $k = 1, \ldots, N$.

At this point, summing over k and dividing by N both in Eq. (11) and Eq. (13), one trivially obtains bounds for the average parameters $y_{n,a,N} = \sum_{k=1}^{N} Y_{n,a}^{(k)}/N$ and $h_{n,a,N} = \sum_{k=1}^{N} H_{n,a}^{(k)}/N$ from the round-dependent bounds. Namely, defining $\overline{Z}_{a,N} = Z_{a,N}/N$ and $\overline{E}_{a,N} = E_{a,N}/N$, the final bounds are

$$\begin{aligned} \frac{\left\langle \overline{Z}_{a,N} \right\rangle}{q_{\mathrm{Z}}^2 p_a} &\geq e^{-a^+} y_{0,a} + \sum_{n=1}^{n_{\rm cut}} \frac{e^{-a^-} a^{-\,n}}{n!} y_{n,a,N}, \\ \frac{\left\langle \overline{Z}_{a,N} \right\rangle}{q_{\mathrm{Z}}^2 p_a} &\leq 1 - e^{-a^+} + e^{-a^-} y_{0,a} - \sum_{n=1}^{n_{\rm cut}} \frac{e^{-a^+} a^{+\,n}}{n!} \left(1 - y_{n,a,N}\right) \quad \text{and} \\ \frac{\left\langle \overline{E}_{a,N} \right\rangle}{q_{\mathrm{X}}^2 p_a} &\geq e^{-a^+} h_{0,a} + \sum_{n=1}^{n_{\rm cut}} \frac{e^{-a^-} a^{-\,n}}{n!} h_{n,a,N}, \\ \frac{\left\langle \overline{E}_{a,N} \right\rangle}{q_{\mathrm{X}}^2 p_a} &\leq 1 - e^{-a^+} + e^{-a^-} h_{0,a} - \sum_{n=1}^{n_{\rm cut}} \frac{e^{-a^+} a^{+\,n}}{n!} \left(1 - h_{n,a,N}\right), \end{aligned} \quad (14)$$

for a common threshold photon number $n_{\rm cut}$ and for all $a \in A$.

2.4 Linear programs for parameter estimation

Even though Eq. (4) provides the relevant restrictions on the maximum bias that Eve can induce between different yields/error probabilities, it consists of a set of non-linear constraints unsuitable for parameter estimation via linear programming. As mentioned in Sec. 2.2 though, in virtue of the convexity/concavity of the functions that define the constraints, their first-order expansions around any given reference yield/error provide valid linear bounds as well. For instance, if we focus on the yields, we have that $G_-(Y_{n,a}^{(k)}, \tau_{ab,n}^{\xi}) \geq G_-(\tilde{Y}_{n,a}^{(k)}, \tau_{ab,n}^{\xi}) + G'_-(\tilde{Y}_{n,a}^{(k)}, \tau_{ab,n})(Y_{n,a}^{(k)} - \tilde{Y}_{n,a}^{(k)})$ and $G_+(Y_{n,a}^{(k)}, \tau_{ab,n}^{\xi}) \leq G_+(\tilde{Y}_{n,a}^{(k)}, \tau_{ab,n}^{\xi}) +$

$G'_+(\tilde{Y}^{(k)}_{n,a}, \tau^{\xi}_{ab,n})(Y^{(k)}_{n,a} - \tilde{Y}^{(k)}_{n,a})$ for all $Y^{(k)}_{n,a} \in (0,1)$, irrespectively of which reference yields $\tilde{Y}^{(k)}_{n,a} \in (0,1)$ we select for the linear expansion. Also, note that the derivative functions $G'_{\pm}$ are well-defined for all $Y^{(k)}_{n,a} \in (0,1)$ because the $G_{\pm}$ are smooth piecewise functions. In particular,

$$G'_-(y,z) = \begin{cases} g'_-(y,z) & \text{if} \quad y > 1-z \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad G'_+(y,z) = \begin{cases} g'_+(y,z) & \text{if} \quad y < z \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

with $g'_{\pm}(y,z) = -1 + 2z \pm (1-2y)\sqrt{z(1-z)/y(1-y)}$. Thus, given a reference yield $\tilde{Y}^{(k)}_{n,a} \in (0,1)$, the linearized bounds are

$$\begin{aligned} &G_-\left(\tilde{Y}^{(k)}_{n,a}, \tau^{\xi}_{ab,n}\right) + G'_-\left(\tilde{Y}^{(k)}_{n,a}, \tau^{\xi}_{ab,n}\right)\left(Y^{(k)}_{n,a} - \tilde{Y}^{(k)}_{n,a}\right) \leq Y^{(k)}_{n,b} \leq \\ &G_+\left(\tilde{Y}^{(k)}_{n,a}, \tau^{\xi}_{ab,n}\right) + G'_+\left(\tilde{Y}^{(k)}_{n,a}, \tau^{\xi}_{ab,n}\right)\left(Y^{(k)}_{n,a} - \tilde{Y}^{(k)}_{n,a}\right). \end{aligned} \tag{16}$$

Identically, for any given reference n-photon error click probabilities $\tilde{H}^{(k)}_{n,a,r} \in (0,1)$, the linearized versions of the corresponding constraints read

$$\begin{aligned} &G_-\left(\tilde{H}^{(k)}_{n,a,r}, \tau^{\xi}_{ab,n}\right) + G'_-\left(\tilde{H}^{(k)}_{n,a,r}, \tau^{\xi}_{ab,n}\right)\left(H^{(k)}_{n,a,r} - \tilde{H}^{(k)}_{n,a,r}\right) \leq H^{(k)}_{n,b,r} \leq \\ &G_+\left(\tilde{H}^{(k)}_{n,a,r}, \tau^{\xi}_{ab,n}\right) + G'_+\left(\tilde{H}^{(k)}_{n,a,r}, \tau^{\xi}_{ab,n}\right)\left(H^{(k)}_{n,a,r} - \tilde{H}^{(k)}_{n,a,r}\right). \end{aligned} \tag{17}$$

If, in addition, we select reference parameters independent of r, say $\tilde{H}^{(k)}_{n,a,r} = \tilde{H}^{(k)}_{n,a}$ for both $r = 0$ and $r = 1$, the linearized lower (upper) bound has the exact same slope and the exact same intercept for both $r = 0$ and $r = 1$. As a consequence, the relevant error probabilities entering the decoy-state analysis, $H^{(k)}_{n,a} = (H^{(k)}_{n,a,0} + H^{(k)}_{n,a,1})/2$ and $H^{(k)}_{n,b} = (H^{(k)}_{n,b,0} + H^{(k)}_{n,b,1})/2$, trivially verify

$$\begin{aligned} &G_-\left(\tilde{H}^{(k)}_{n,a}, \tau^{\xi}_{ab,n}\right) + G'_-\left(\tilde{H}^{(k)}_{n,a}, \tau^{\xi}_{ab,n}\right)\left(H^{(k)}_{n,a} - \tilde{H}^{(k)}_{n,a}\right) \leq H^{(k)}_{n,b} \leq \\ &G_+\left(\tilde{H}^{(k)}_{n,a}, \tau^{\xi}_{ab,n}\right) + G'_+\left(\tilde{H}^{(k)}_{n,a}, \tau^{\xi}_{ab,n}\right)\left(H^{(k)}_{n,a} - \tilde{H}^{(k)}_{n,a}\right), \end{aligned} \tag{18}$$

as we wanted to show.

As a final comment, note that, for all practical purposes, one can restrict the reference parameters to be round-independent: $\tilde{Y}^{(k)}_{n,a} = \tilde{y}_{n,a}$ and $\tilde{H}^{(k)}_{n,a} = \tilde{h}_{n,a}$ for all $k = 1, \ldots, N$. This being the case, summing over k and dividing by N in Eq. (16) and Eq. (18), one obtains respective inequalities for the average parameters $y_{n,b,N} = \sum_{k=1}^{N} Y^{(k)}_{n,b}/N$ and $h_{n,b,N} = \sum_{k=1}^{N} H^{(k)}_{n,b}/N$. Namely, for all $a \in A$, $b \in A$ $(b \neq a)$ and $n \in \mathbb{N}$, we have $c^{-}_{ab,n} + m^{-}_{ab,n} y_{n,a,N} \leq y_{n,b,N} \leq c^{+}_{ab,n} + m^{+}_{ab,n} y_{n,a,N}$ and $t^{-}_{ab,n} + s^{-}_{ab,n} h_{n,a,N} \leq h_{n,b,N} \leq t^{+}_{ab,n} + s^{+}_{ab,n} h_{n,a,N}$, where, for conciseness, we define the intercepts and slopes

$$\begin{aligned} &c^{\pm}_{ab,n} = G_{\pm}(\tilde{y}_{n,a}, \tau^{\xi}_{ab,n}) - G'_{\pm}(\tilde{y}_{n,a}, \tau^{\xi}_{ab,n})\tilde{y}_{n,a}, \quad m^{\pm}_{ab,n} = G'_{\pm}(\tilde{y}_{n,a}, \tau^{\xi}_{ab,n}), \\ &t^{\pm}_{ab,n} = G_{\pm}(\tilde{h}_{n,a}, \tau^{\xi}_{ab,n}) - G'_{\pm}(\tilde{h}_{n,a}, \tau^{\xi}_{ab,n})\tilde{h}_{n,a} \ \text{ and } \ s^{\pm}_{ab,n} = G'_{\pm}(\tilde{h}_{n,a}, \tau^{\xi}_{ab,n}). \end{aligned} \tag{19}$$

Of course, the tightness of these linear bounds is subject to the adequacy of the selected reference yields, and thus it relies on a characterization of the quantum channel. Note, however, that aiming to further improve the results, one could incorporate more linearized CS constraints to the problem by using various reference yields for each pair (n, a), instead of just one. Also, we recall that simpler bounds not relying on any reference values can

be derived by using the so-called trace distance argument [27], and this is what we do in Appendix B. Another alternative would be, of course, to solve a non-linear optimization problem given by the original CS constraints.

To finish with, we present the linear programs that allow to estimate the relevant single-photon parameters, putting together the decoy-state constraints introduced in Sec. 2.3 and the above linearized CS constraints. In the first place, we address the average number of signal-setting single-photon counts, defined as $\overline{Z}_{1,\mu,N} = \sum_{k=1}^{N} Z_{1,\mu}^{(k)}/N$ with $Z_{1,\mu}^{(k)} = Z_{\mu}^{(k)} \mathbb{1}_{\{n_k=1\}}$. Since $\langle Z_{1,\mu}^{(k)} \rangle = q_{\mathrm{Z}}^2 p_\mu p^{(k)}(1|\mu) Y_{1,\mu}^{(k)} \geq q_{\mathrm{Z}}^2 p_\mu \mu^- e^{-\mu^-} Y_{1,\mu}^{(k)}$, averaging over k it follows that

$$\langle \overline{Z}_{1,\mu,N} \rangle \geq q_{\mathrm{Z}}^2 p_\mu \mu^- e^{-\mu^-} y_{1,\mu,N}, \tag{20}$$

and a lower bound $y_{1,\mu,N}^{\mathrm{L}}$ on $y_{1,\mu,N}$ is reached by the following linear program:

$$\begin{aligned}
&\min \quad y_{1,\mu,N} \qquad\qquad (21)\\
\text{s.t.} \quad &\frac{\left\langle \overline{Z}_{a,N} \right\rangle}{q_{\mathrm{Z}}^2 p_a} \geq e^{-a^+} y_{0,a} + \sum_{n=1}^{n_{\mathrm{cut}}} \frac{e^{-a^-} a^{-\,n}}{n!} y_{n,a,N} \quad (a \in A),\\
&\frac{\left\langle \overline{Z}_{a,N} \right\rangle}{q_{\mathrm{Z}}^2 p_a} \leq 1 - e^{-a^+} + e^{-a^-} y_{0,a} - \sum_{n=1}^{n_{\mathrm{cut}}} \frac{e^{-a^+} a^{+\,n}}{n!} (1 - y_{n,a,N}) \quad (a \in A),\\
&c_{ab,n}^{+} + m_{ab,n}^{+} y_{n,a,N} \geq y_{n,b,N} \quad (a \in A,\ b \in A,\ b \neq a,\ n = 0, \ldots, n_{\mathrm{cut}}),\\
&c_{ab,n}^{-} + m_{ab,n}^{-} y_{n,a,N} \leq y_{n,b,N} \quad (a \in A,\ b \in A,\ b \neq a,\ n = 0, \ldots, n_{\mathrm{cut}}),\\
&0 \leq y_{n,a,N} \leq 1 \quad (a \in A, n = 0, \ldots, n_{\mathrm{cut}}).
\end{aligned}$$

We recall that the $c_{ab,n}^{\pm}$ and the $m_{ab,n}^{\pm}$ are defined in Eq. (19). Needless to say, replacing Z by X everywhere one obtains the corresponding program for the average number of signal-setting single-photon counts in the X basis, $\overline{X}_{1,\mu,N}$, such that

$$\langle \overline{X}_{1,\mu,N} \rangle \geq q_{\mathrm{X}}^2 p_\mu \mu^- e^{-\mu^-} y'_{1,\mu,N}, \tag{22}$$

where the apostrophe here denotes that we refer to the X basis.

On the other hand, the average number of signal-setting single-photon error counts in the X basis is $\overline{E}_{1,\mu,N} = \sum_{k=1}^{N} E_{1,\mu}^{(k)}/N$, with $E_{1,\mu}^{(k)} = E_{\mu}^{(k)} \mathbb{1}_{\{n_k=1\}}$. Since $\langle E_{1,\mu}^{(k)} \rangle = q_{\mathrm{X}}^2 p_\mu p^{(k)}(1|\mu) H_{1,\mu}^{(k)} \leq q_{\mathrm{X}}^2 p_\mu \mu^+ e^{-\mu^+} H_{1,\mu}^{(k)}$, averaging over k it follows that

$$\langle \overline{E}_{1,\mu,N} \rangle \leq q_{\mathrm{X}}^2 p_\mu \mu^+ e^{-\mu^+} h_{1,\mu,N}, \tag{23}$$

and an upper bound $h_{1,\mu,N}^{\mathrm{U}}$ on $h_{1,\mu,N}$ is reached by the following linear program:

$$\begin{aligned}
&\max \quad h_{1,\mu,N} \qquad\qquad (24)\\
\text{s.t.} \quad &\frac{\left\langle \overline{E}_{a,N} \right\rangle}{q_{\mathrm{X}}^2 p_a} \geq e^{-a^+} h_{0,a} + \sum_{n=1}^{n_{\mathrm{cut}}} \frac{e^{-a^-} a^{-\,n}}{n!} h_{n,a,N} \quad (a \in A),\\
&\frac{\left\langle \overline{E}_{a,N} \right\rangle}{q_{\mathrm{X}}^2 p_a} \leq 1 - e^{-a^+} + e^{-a^-} h_{0,a} - \sum_{n=1}^{n_{\mathrm{cut}}} \frac{e^{-a^+} a^{+\,n}}{n!} (1 - h_{n,a,N}) \quad (a \in A),\\
&t_{ab,n}^{+} + s_{ab,n}^{+} h_{n,a,N} \geq h_{n,b,N} \quad (a \in A,\ b \in A,\ a \neq b,\ n = 0, \ldots, n_{\mathrm{cut}}),\\
&t_{ab,n}^{-} + s_{ab,n}^{-} h_{n,a,N} \leq h_{n,b,N} \quad (a \in A,\ b \in A,\ a \neq b,\ n = 0, \ldots, n_{\mathrm{cut}}),\\
&0 \leq h_{n,a,N} \leq 1 \quad (a \in A, n = 0, \ldots, n_{\mathrm{cut}}).
\end{aligned}$$

Finally, we remark that, in virtue of the properties of linear optimization [29], $y^{\mathrm{L}}_{1,\mu,N}$, ${y'}^{\mathrm{L}}_{1,\mu,N}$ and $h^{\mathrm{U}}_{1,\mu,N}$ are linear in $\left\langle \overline{Z}_{a,N} \right\rangle$, $\left\langle \overline{X}_{a,N} \right\rangle$ and $\left\langle \overline{E}_{a,N} \right\rangle$, respectively, for all $a \in A$, which means that they provide the expectation of certain random variables respectively linear in $\overline{Z}_{a,N}$, $\overline{X}_{a,N}$ and $\overline{E}_{a,N}$. In turn, this implies that the bounds reached by the linear programs can be written as $\langle \overline{Z}_{1,\mu,N} \rangle \geq \langle \overline{Z}^{\mathrm{L}}_{1,\mu,N} \rangle$, $\langle \overline{X}_{1,\mu,N} \rangle \geq \langle \overline{X}^{\mathrm{L}}_{1,\mu,N} \rangle$ and $\langle \overline{E}_{1,\mu,N} \rangle \leq \langle \overline{E}^{\mathrm{U}}_{1,\mu,N} \rangle$, where

$$\begin{aligned} \langle \overline{Z}^{\mathrm{L}}_{1,\mu,N} \rangle &= q_{\mathrm{Z}}^2 p_\mu \mu^- e^{-\mu^-} y^{\mathrm{L}}_{1,\mu,N}, \\ \langle \overline{X}^{\mathrm{L}}_{1,\mu,N} \rangle &= q_{\mathrm{X}}^2 p_\mu \mu^- e^{-\mu^-} {y'}^{\mathrm{L}}_{1,\mu,N} \quad \text{and} \\ \langle \overline{E}^{\mathrm{U}}_{1,\mu,N} \rangle &= q_{\mathrm{X}}^2 p_\mu \mu^+ e^{-\mu^+} h^{\mathrm{U}}_{1,\mu,N} \end{aligned} \tag{25}$$

for all N. This feature is crucial to justify the asymptotic approximation of the secret key rate presented next.

2.5 Asymptotic approximation of the secret key rate

The linear programs of Sec. 2.4 provide suitable bounds on the expectations of the relevant experimental averages, namely, the average number of signal-setting single-photon counts in the Z (X) basis after N transmission rounds, $\overline{Z}_{1,\mu,N}$ ($\overline{X}_{1,\mu,N}$), and the average number of signal-setting single-photon error counts in the X basis after N transmission rounds, $\overline{E}_{1,\mu,N}$. The bounds are of the form

$$\langle \overline{Z}_{1,\mu,N} \rangle \geq \langle \overline{Z}^{\mathrm{L}}_{1,\mu,N} \rangle, \;\; \langle \overline{X}_{1,\mu,N} \rangle \geq \langle \overline{X}^{\mathrm{L}}_{1,\mu,N} \rangle \;\text{ and }\; \langle \overline{E}_{1,\mu,N} \rangle \leq \langle \overline{E}^{\mathrm{U}}_{1,\mu,N} \rangle \tag{26}$$

for all N, where $\overline{Z}^{\mathrm{L}}_{1,\mu,N}$ ($\overline{X}^{\mathrm{L}}_{1,\mu,N}$) is a linear combination of the experimentally observed Z (X) basis gains, $\{\overline{Z}_{a,N}\}_{a\in A}$ ($\{\overline{X}_{a,N}\}_{a\in A}$), and $\overline{E}^{\mathrm{U}}_{1,\mu,N}$ is a linear combination of the experimentally observed numbers of errors in the X basis, $\{\overline{E}_{a,N}\}_{a\in A}$.

However, the finite secret key rate relies on statistical bounds of the experimental averages themselves, say

$$P\left(\overline{Z}_{1,\mu,N} < \overline{Z}^{\mathrm{L},\epsilon_1}_{1,\mu,N}\right) \leq \epsilon_1, \;\; P\left(\overline{X}_{1,\mu,N} < \overline{X}^{\mathrm{L},\epsilon_2}_{1,\mu,N}\right) \leq \epsilon_2 \;\text{ and }\; P\left(\overline{E}_{1,\mu,N} > \overline{E}^{\mathrm{U},\epsilon_3}_{1,\mu,N}\right) \leq \epsilon_3 \tag{27}$$

for given failure probabilities ϵ_1, ϵ_2 and ϵ_3 (see the Methods Sec. 4.2 for a summarized derivation of the finite secret key rate). Crucially, in the absence of intensity correlations or side-channels possibly leaking the intensity setting information, commutativity allows to consider the so-called counterfactual setting, in which case the latter bounds —Eq. (27)— are obtained from the former —Eq. (26)— via concentration inequalities for independent random variables, such as Chernoff's [24] or Hoeffding's [25]. Precisely, the independence of the relevant indicator variables attached to the detection events is enforced because, in the counterfactual setting, the intensities are randomly selected a posteriori (and thus decoupled) of the detection events. On the contrary, intensity correlations invalidate the counterfactual setting argument, in so invalidating the usage of the above concentration inequalities too. Nevertheless, propositions 1 and 2 in the Methods Sec. 4.3 establish that, as long as the variance of the experimental averages tends to zero as N tends to infinity, any violation of the equations

$$\overline{Z}_{1,\mu,N} \geq \overline{Z}^{\mathrm{L}}_{1,\mu,N}, \;\; \overline{X}_{1,\mu,N} \geq \overline{X}^{\mathrm{L}}_{1,\mu,N} \;\text{ and }\; \overline{E}_{1,\mu,N} \leq \overline{E}^{\mathrm{U}}_{1,\mu,N} \tag{28}$$

—no matter how small— has an asymptotically null probability of occurring.

This feature legitimizes the use of the bounds of Eq. (28) to asymptotically approximate the secret key rate, by plugging them into the finite secret key rate formula —Eq. (61) in the Methods Sec. 4.2—. This yields

$$K_N \approx \overline{Z}^{\mathrm{L}}_{1,\mu,N}\left[1-h\left(\frac{\overline{E}^{\mathrm{U}}_{1,\mu,N}}{\overline{X}^{\mathrm{L}}_{1,\mu,N}}+\sqrt{\frac{\left(\overline{X}^{\mathrm{L}}_{1,\mu,N}+\overline{Z}^{\mathrm{L}}_{1,\mu,N}\right)\left(\overline{Z}^{\mathrm{L}}_{1,\mu,N}+1/N\right)}{2N\overline{Z}^{\mathrm{L}\,2}_{1,\mu,N}\overline{X}^{\mathrm{L}}_{1,\mu,N}}\log\left(\frac{1}{\epsilon_{\mathrm{S}}}\right)}\right)\right]$$
$$-f_{\mathrm{EC}}\overline{Z}_{\mu,N}h(E_{\mathrm{tol}})-\frac{1}{N}\log\left(\frac{1}{\epsilon_{\mathrm{cor}}\epsilon^2_{\mathrm{PA}}\delta}\right), \tag{29}$$

for large enough N, with secrecy parameter $\epsilon_{\mathrm{sec}} \approx 2\epsilon_{\mathrm{S}}+\epsilon_{\mathrm{PA}}+\delta$ and correctness parameter ϵ_{cor} (see the Methods Sec. 4.2 for the definitions of the security parameters ϵ_{S}, ϵ_{PA} and δ). Also, $h(\cdot)$ denotes the binary entropy function, f_{EC} is the efficiency of the error correction protocol and E_{tol} is a threshold bit error rate. Note that Eq. (29) assumes that Alice and Bob use the Z (X) basis events for key generation (parameter estimation).

Having reached this stage, one can remove both the dependence on N and on the security parameters by neglecting the Serfling deviation term (which scales as $N^{-1/2}$) and the finite key term $\log\left(1/\epsilon_{\mathrm{cor}}\epsilon^2_{\mathrm{PA}}\delta\right)/N$, in so obtaining the final asymptotic secret key rate formula

$$K_\infty = \overline{Z}^{\mathrm{L}}_{1,\mu,N}\left[1-h\left(\frac{\overline{E}^{\mathrm{U}}_{1,\mu,N}}{\overline{X}^{\mathrm{L}}_{1,\mu,N}}\right)\right]-f_{\mathrm{EC}}\overline{Z}_{\mu,N}h(E_{\mathrm{tol}}). \tag{30}$$

Notably, as mentioned, the usefulness of this asymptotic approximation is subject to the non-trivial condition that the variance of the experimental averages vanishes as N tends to infinity. Precisely, for a sequence $\{X_j\}$ with successive averages $\overline{X}_N = \sum_{j=1}^N X_j/N$, we have that $Var[\overline{X}_N] = \sum_{i=1}^N Var[X_i]/N^2 + 2\sum_{i=1}^N\sum_{j>i}^N Cov[X_i,X_j]/N^2$. Then, since $\lim_{N\to\infty}\sum_{i=1}^N Var[X_i]/N^2 \leq \lim_{N\to\infty}\max_i\{Var[X_i]\}/N = 0$, it follows that $\lim_{N\to\infty} Var[\overline{X}_N] = 2\times\lim_{N\to\infty}\sum_{i=1}^N\sum_{j>i}^N Cov[X_i,X_j]/N^2$, as long as the latter is finite. Thus, in particular, Eq. (30) is an asymptotic approximation of the secret key rate provided that

$$\lim_{N\to\infty}\sum_{i=1}^N\sum_{j>i}^N\frac{Cov[X_i,X_j]}{N^2}=0 \tag{31}$$

for the relevant sequences $X_k \in \{Z^{(k)}_a, X^{(k)}_a, E^{(k)}_a, Z^{(k)}_{1,a}, X^{(k)}_{1,a}, E^{(k)}_{1,a}\}_{a\in A}$ (*i.e.,* provided that the preconditions of propositions 1 and 2 in the Methods Sec. 4.3 hold). For instance, if $Cov\left[X_i,X_j\right]=0$ for all i, j with $|i-j|>\zeta$ —where ζ denotes a finite round difference— the condition holds despite the fact that the X_j may be dependent and non-identically distributed.

2.6 Simulations

In the absence of real data, we fix the experimental inputs $\overline{Z}_{a,N}/q^2_Z p_a$, $\overline{X}_{a,N}/q^2_X p_a$ and $\overline{E}_{a,N}/q^2_X p_a$ of the linear programs to their expected values according to a typical channel model. Importantly, although the security analysis contemplates intensity correlations, we adopt a standard channel and transmitter model without correlations for ease of comparison with prior work. In particular, let η_{det} ($\eta_{\mathrm{ch}} = 10^{-\alpha_{\mathrm{att}}L/10}$) denote the detection efficiency of Bob's detectors (transmittance of the channel), where α_{att} (dB/km) is the attenuation coefficient of the channel and L (km) is the distance between Alice's and Bob's labs.

Also, let p_{d} (δ_{A}) stand for the dark count probability of each of Bob's photo-detectors (polarization misalignment occurring in the channel). The model is [15, 26]

$$\begin{aligned}
&\frac{\left\langle \overline{Z}_{a,N} \right\rangle}{q_{\mathrm{Z}}^2 p_a} = \frac{\left\langle \overline{X}_{a,N} \right\rangle}{q_{\mathrm{X}}^2 p_a} = 1-(1-p_{\mathrm{d}})^2 e^{-\eta a} \text{ and} \\
&\frac{\left\langle \overline{E}_{a,N} \right\rangle}{q_{\mathrm{X}}^2 p_a} = \frac{\left\langle \overline{E}_{a,N(\mathrm{Z})} \right\rangle}{q_{\mathrm{Z}}^2 p_a} = \frac{p_{\mathrm{d}}^2}{2} + p_{\mathrm{d}}(1-p_{\mathrm{d}})(1+h_{\eta,a,\delta_{\mathrm{A}}}) \\
&+(1-p_{\mathrm{d}})^2 \left(\frac{1}{2} + h_{\eta,a,\delta_{\mathrm{A}}} - \frac{1}{2} e^{-\eta a} \right)
\end{aligned} \tag{32}$$

for $a \in A$, where $\eta = \eta_{\mathrm{det}}\eta_{\mathrm{ch}}$ and we define $h_{\eta,a,\delta_{\mathrm{A}}} = \left(e^{-\eta a \cos^2 \delta_{\mathrm{A}}} - e^{-\eta a \sin^2 \delta_{\mathrm{A}}}\right)/2$. Also, we introduce the variable $\overline{E}_{a,N(\mathrm{Z})}$, which is equivalent to $\overline{E}_{a,N}$ but referred to the Z basis error clicks. Note that Eq. (32) accounts for the fact that multiple clicks are randomly assigned to a specific detection outcome. For simplicity, the tolerated bit error rate of the sifted key is set to $E_{\mathrm{tol}} = \langle \overline{E}_{\mu,N(\mathrm{Z})} \rangle / \langle \overline{Z}_{\mu,N} \rangle$.

In addition, we remark that the channel model can also be used to select reference values $\tilde{y}_{n,a,N}$ and $\tilde{h}_{n,a,N}$ for the evaluation of the linearized CS constraints of Sec. 2.4. Nevertheless, one could also choose the reference values based on previous executions of the protocol instead. Note that, in a real QKD experiment, these reference values would be required for the parameter estimation to go through, and so the secret key rate would be sensitive to the selected $\tilde{y}_{n,a,N}$ and $\tilde{h}_{n,a,N}$. The reader is referred to Appendix A for the explicit formulas of $\tilde{y}_{n,a,N}$ and $\tilde{h}_{n,a,N}$ that we use, matching the typical channel model under consideration. Alternatively, in Appendix B we provide a looser analysis based on the trace distance argument [27], which does not rely on the selection of any reference values.

Either way, plugging Eq. (32) into Eq. (21) and Eq. (24), we find that the asymptotic secret key rate formula K_∞ (presented in Eq. (30)) does not depend either on the probability of the decoy settings, p_ν and p_ω, or on the probability of selecting the X basis, q_{X}, in such a way that setting $p_\mu \approx 1$ and $q_{\mathrm{Z}} \approx 1$ maximizes K_∞ with the typical channel model under consideration. This feature corroborates the intuition that, as N increases, one can devote larger and larger fractions of rounds to key generation without compromising the tightness of the parameter estimation. Lastly, regarding the experimental parameters, we list them below. We take $\eta_{\mathrm{det}} = 0.65$, $p_{\mathrm{d}} = 7.2 \times 10^{-8}$ —both values matching the recent experiment reported in [6]—, a typical attenuation coefficient $\alpha_{\mathrm{att}} = 0.2$ dB/km and a standard error correction efficiency of $f_{\mathrm{EC}} = 1.16$. Regarding the misalignment, we take $\delta_{\mathrm{A}} = 0.08$ for illustration purposes. Also, we fix the weakest intensity setting to $\omega = 10^{-4}$ for the numerics, and we numerically optimize μ and ν to maximize K_∞ for each value of the distance L. Lastly, three different correlation ranges are contemplated, $\xi = 1$, $\xi = 2$ and $\xi = 5$, each of which is combined with various values of the maximum relative deviation δ_{max} (see assumptions 2 and 3 in Sec. 2.1 for the definitions of ξ and δ_{max}).

The rate-distance performance with the above considerations is shown in Fig. 1. As seen in the figure, intensity correlations strongly limit the maximum distance attainable for QKD, and the secret key rate is notably sensitive to the deviation parameter, δ_{max}. On the contrary, as long as moderate values are considered for the correlation range ξ, the effect of this parameter on the secret key rate is softer. Finally, for completeness, in Appendix C we show that an enhancement of the secret key rate is possible by assuming deterministic intensity correlations, in contrast to the model-independent correlations considered so far.

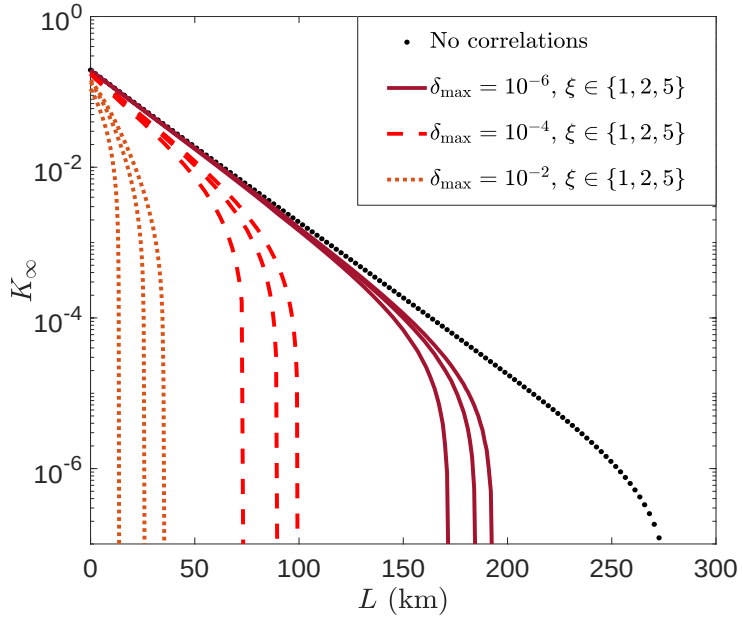

Figure 1: QKD performance in the presence of finite range intensity correlations. The figure shows the asymptotic secret key rate, K_∞, as a function of the distance, L, when performing the parameter estimation with the linearized CS constraint. For illustration purposes, we consider three different values of the maximum relative deviation between intensity settings (a_k) and actual intensities (α_k), $\delta_{\max} \in \{10^{-6}, 10^{-4}, 10^{-2}\}$, and three correlation ranges are contemplated in each case, $\xi \in \{1, 2, 5\}$. Moreover, for comparison purposes, we show the attainable secret key rate in the absence of intensity correlations as well (dotted black line). The experimental parameters are fixed as specified in the main text.

3 Discussion

Aiming to enhance the performance of QKD systems, it is crucial to develop ultrafast clock rate QKD devices capable of delivering high secret key rates for widespread applications. However, even for GHz clock rates, QKD transmitters exhibit intensity correlations [11, 10] that invalidate standard decoy-state analyses for parameter estimation. As opposed to many other source imperfections, only limited solutions were known for this security loophole so far [10, 17, 18]. In this work, we solve the problem by quantifying the maximum effect of intensity correlations in the security of QKD. For this purpose, we introduce two experimental-friendly security parameters that allow to characterize arbitrary correlations in the IM. Importantly, our technique builds on a result that we refer to as the Cauchy-Schwarz constraint (recently used in [20, 28]), which provides tighter bounds on the indistinguishability of non-orthogonal quantum states than the well-known trace distance argument [27].

For illustration purposes, our analysis is dedicated to the standard decoy-state BB84 protocol with one signal and two decoy settings [15], although we remark that our results can be easily generalized to deal with other variants of the protocol, or even with the decoy-state measurement-device-independent (MDI) QKD scheme [21].

As a related contribution, we present a non-standard derivation of the asymptotic limit, in so revealing a necessary condition to justify the ubiquitous asymptotic formula. Crucially, this condition becomes non-trivial in the most general context of coherent attacks and arbitrary pulse correlations. Nevertheless, if, for instance, Eve's attack does not interrelate arbitrarily distant detection events —but the interaction is limited to a finite round

difference— the condition holds. In this regard, it is not unreasonable to conjecture that, as long as the correlation range of the light pulses is not larger than ξ, Eve may reach an optimal cheating strategy by attacking blocks of maximum round difference $\zeta = \xi$. Indeed, a hypothesis of this kind might pave the way for a finite key analysis of the problem, which is the natural direction to follow. In any case, whether or not this conjecture is true, the solution here presented clearly provides an insightful step towards foolproof security of high speed QKD systems, with their imperfections.

4 Methods

4.1 Cauchy-Schwarz constraint

The CS constraint is stated as follows [20, 23].

Theorem [20]. Let $|u\rangle$ and $|v\rangle$ be pure states of a certain quantum system. Then, for all positive operators $\hat{O} \leq I$,

$$G_{-}\left(\langle u|\hat{O}|u\rangle, |\langle v|u\rangle|^2\right) \leq \langle v|\hat{O}|v\rangle \leq G_{+}\left(\langle u|\hat{O}|u\rangle, |\langle v|u\rangle|^2\right), \tag{33}$$

where the functions $G_{\pm}$ are defined in Eq. (5). As pointed out in Sec. (2.2), for all $k = 1, \ldots, N$, all $n \in \mathbb{N}$, and any given pair of settings, $a \in A$ and $b \in A$ with $b \neq a$, Eq. (33) allows to constrain the maximum deviation that Eve can induce between the n-photon yields $p^{(k)}(\text{click}|n, a, \text{Z}, \text{Z}) = p^{(k)}(\text{click}|n, a, \text{Z})$ and $p^{(k)}(\text{click}|n, b, \text{Z}, \text{Z}) = p^{(k)}(\text{click}|n, b, \text{Z})$, where we have invoked the basis-independent detection efficiency assumption to remove the conditioning on Bob's basis choice. Here, we derive the specific constraint, namely, Eq. (4), which contemplates fully general coherent attacks and finite range intensity correlations.

In an entanglement based view of the protocol, the global input state describing all the protocol rounds reads

$$|\Psi\rangle = \left[\sum_{a_1^N}\sum_{x_1^N}\sum_{r_1^N}\left(\prod_{i=1}^{N}\sqrt{\frac{p_{a_i}q_{x_i}}{2}}\right)\left(\bigotimes_{i=1}^{N}|a_i, x_i\rangle_{A_i}|r_i\rangle_{A_i'}\left|\psi_{\vec{a}_i}^{x_i,r_i}\right\rangle_{B_iC_i}\right)\right]\otimes|0\rangle_E, \tag{34}$$

where we introduce the notation $a_1^N = a_1 \ldots a_N$, and equivalently for x_1^N and r_1^N. Also, for all i, $\{|a_i, x_i\rangle_{A_i}|a_i \in A, x_i \in M\}$ and $\{|r_i\rangle_{A_i'}|r_i \in \mathbb{Z}_2\}$ are orthonormal bases of Alice's i-th registers, A_i and $A_i^{'}$. Similarly, we define

$$\left|\psi_{\vec{a}_i}^{x_i,r_i}\right\rangle_{B_iC_i} = \sum_{n_i=0}^{\infty}\sqrt{p_{n_i|\vec{a}_i}}\,|t_{n_i}\rangle_{C_i}|n_i^{x_i,r_i}\rangle_{B_i}, \tag{35}$$

where C_i denotes an inaccessible purifying system with orthonormal basis $\{|t_{n_i}\rangle_{C_i}|n_i \in \mathbb{N}\}$ for all i (C_i stores the photon number information of the i-th signal that Alice sends to Bob), B_i denotes the system delivered to Bob ($|n_i^{x_i,r_i}\rangle_{B_i}$ standing for a Fock state with n_i photons encoding the BB84 polarization state defined by (x_i, r_i)), and the photon number statistics $p_{n_i|\vec{a}_i}$ are defined in Eq. (3). Lastly, $|0\rangle_E$ in Eq. (34) stands for the initial state of Eve's ancillary system.

If we denote Eve's coherent interaction with systems $B_1, \ldots, B_N$ and E by $\hat{U}_{BE}$ — such that $\hat{U}_{BE}|\Psi\rangle$ represents the global state prior to Bob's measurements— and refer to Bob's "click" POVM element in round k as $\hat{M}_{B_k}^{\text{click}} = \mathbb{1}_{B_k} - \hat{M}_{B_k}^{f}$, the joint probability

$p^{(k)}(\text{click}, n, a, \text{Z})$ is computed as

$$\begin{aligned} p^{(k)}(\text{click}, n, a, \text{Z}) &= \text{Tr}\left\{\hat{P}_{|a,\text{Z},t_n\rangle_{A_kC_k}} \hat{M}^{\text{click}}_{B_k} \hat{U}_{BE} \left|\Psi\right\rangle\!\left\langle\Psi\right| \hat{U}^\dagger_{BE}\right\} \\ &= \text{Tr}\left\{\hat{U}^\dagger_{BE} \hat{M}^{\text{click}}_{B_k} \hat{U}_{BE} \hat{P}_{|a,\text{Z},t_n\rangle_{A_kC_k}} \left|\Psi\right\rangle\!\left\langle\Psi\right| \hat{P}_{|a,\text{Z},t_n\rangle_{A_kC_k}}\right\} \\ &= \text{Tr}_{\underline{A_kC_k},A_1'^N B_1^N E}\left\{\hat{U}^\dagger_{BE}\, \hat{M}^{\text{click}}_{B_k}\, \hat{U}_{BE} \left|\widetilde{\Psi}^{(k)}_{a,\text{Z},n}\right\rangle\!\left\langle\widetilde{\Psi}^{(k)}_{a,\text{Z},n}\right|\right\}, \end{aligned} \tag{36}$$

where $\hat{P}_{|a,\text{Z},t_n\rangle_{A_kC_k}} = |a,\text{Z}\rangle\langle a,\text{Z}|_{A_k} \otimes |t_n\rangle\langle t_n|_{C_k}$, $\underline{A_kC_k} = \{A_jC_j | j \neq k\}$, and we have introduced the unnormalized pure state

$$\left|\widetilde{\Psi}^{(k)}_{a,\text{Z},n}\right\rangle = \langle a,\text{Z}|_{A_k} \langle t_n|_{C_k} \left|\Psi\right\rangle. \tag{37}$$

Note that, in the derivation of Eq. (36), we make use of the fact that projection operators are "self-squared", together with the cyclic property of the trace and straightforward commutation relations. Then, we trace out systems A_k and C_k explicitly, in order to obtain the necessary input of the CS constraint later on.

Further defining $|\Psi^{(k)}_{a,\text{Z},n}\rangle = |\widetilde{\Psi}^{(k)}_{a,\text{Z},n}\rangle / \| |\widetilde{\Psi}^{(k)}_{a,\text{Z},n}\rangle \|$, Eq. (36) can be restated as

$$p^{(k)}(\text{click}, n, a, \text{Z}) = \left\|\left|\widetilde{\Psi}^{(k)}_{a,\text{Z},n}\right\rangle\right\|^2 \text{Tr}\left\{\hat{U}^\dagger_{BE}\, \hat{M}^{\text{click}}_{B_k}\, \hat{U}_{BE} \left|\Psi^{(k)}_{a,\text{Z},n}\right\rangle\!\left\langle\Psi^{(k)}_{a,\text{Z},n}\right|\right\}, \tag{38}$$

and since $p^{(k)}(n, a, \text{Z}) = \text{Tr}\left\{\hat{P}_{|a,\text{Z},t_n\rangle_{A_kC_k}} \hat{U}_{BE} \left|\Psi\right\rangle\!\left\langle\Psi\right| \hat{U}^\dagger_{BE}\right\} = \left\|\left|\widetilde{\Psi}^{(k)}_{a,\text{Z},n}\right\rangle\right\|^2$, it follows from Bayes rule that

$$p^{(k)}(\text{click}|n, a, \text{Z}) = \text{Tr}\left\{\hat{O}^{(k)}_{\text{click}} \left|\Psi^{(k)}_{a,\text{Z},n}\right\rangle\!\left\langle\Psi^{(k)}_{a,\text{Z},n}\right|\right\} = \left\langle\Psi^{(k)}_{a,\text{Z},n}\right| \hat{O}^{(k)}_{\text{click}} \left|\Psi^{(k)}_{a,\text{Z},n}\right\rangle, \tag{39}$$

where $\hat{O}^{(k)}_{\text{click}} = \hat{U}^\dagger_{BE}\, \hat{M}^{\text{click}}_{B_k}\, \hat{U}_{BE}$. Recalling that $p^{(k)}(\text{click}|n, a, \text{Z}) = p^{(k)}(\text{click}|n, a, \text{Z}, \text{Z}) =: Y^{(k)}_{n,a}$, which is the n-photon yield of round k associated to the intensity setting a, it follows from Eq. (33) that

$$G_-\left(Y^{(k)}_{n,a}, \left|\left\langle\Psi^{(k)}_{b,\text{Z},n}\middle|\Psi^{(k)}_{a,\text{Z},n}\right\rangle\right|^2\right) \leq Y^{(k)}_{n,b} \leq G_+\left(Y^{(k)}_{n,a}, \left|\left\langle\Psi^{(k)}_{b,\text{Z},n}\middle|\Psi^{(k)}_{a,\text{Z},n}\right\rangle\right|^2\right) \tag{40}$$

for all $n \in \mathbb{N}$, $a \in A$, $b \in A$ $(b \neq a)$ and $k = 1, \ldots, N$, and the bounds are tighter the closer $|\langle\Psi^{(k)}_{b,\text{Z},n}|\Psi^{(k)}_{a,\text{Z},n}\rangle|^2$ is to 1. We recall that the interpretation is simple: even if Eve fine-tunes her global unitary $\hat{U}_{BE}$ focusing only in round k, aiming to maximally deviate $Y^{(k)}_{n,a}$ and $Y^{(k)}_{n,b}$, such a deviation is subject to Eq. (40).

In short, evaluating Eq. (40) requires to lower bound $|\langle\Psi^{(k)}_{b,\text{Z},n}|\Psi^{(k)}_{a,\text{Z},n}\rangle|^2$, which we do next. For $k = 2, \ldots, N-1$ (the cases $k = 1$ and $k = N$ will be discussed separately), we have that

$$\begin{aligned} \left|\widetilde{\Psi}^{(k)}_{a,\text{Z},n}\right\rangle = \sqrt{\frac{q_\text{Z} p_a}{2^N}} \Bigg[&\sum_{\underline{a_k}} \sum_{\underline{x_k}} \sum_{r_1^N} \left(\prod_{i\neq k} \sqrt{p_{a_i} q_{x_i}}\right) \left(\bigotimes_{i=1}^{k-1} |a_i, x_i\rangle_{A_i} |r_i\rangle_{A_i'} \left|\psi^{x_i,r_i}_{\vec{a}_i}\right\rangle_{B_iC_i}\right) \times \\ &\left(\sqrt{p_n|_{a,\vec{a}_{k-1}}} |r_k\rangle_{A_k'} \left|n^{\text{Z},r_k}\right\rangle_{B_k}\right) \left(\bigotimes_{i=k+1}^{N} |a_i, x_i\rangle_{A_i} |r_i\rangle_{A_i'} \left|\psi^{x_i,r_i}_{\vec{a}_i(a_k=a)}\right\rangle_{B_iC_i}\right)\Bigg] \otimes |0\rangle_E \end{aligned} \tag{41}$$

with $\underline{a_k} = \{a_j | j \neq k\}$ and $\underline{x_k} = \{x_j | j \neq k\}$. Thus, it follows that

$$\left\langle \widetilde{\Psi}^{(k)}_{b,\mathrm{Z},n} \middle| \widetilde{\Psi}^{(k)}_{a,\mathrm{Z},n} \right\rangle = \frac{q_\mathrm{Z}\sqrt{p_a p_b}}{2^N} \sum_{\underline{a_k}} \sum_{\underline{x_k}} \sum_{r_1^N} \left(\prod_{i\neq k} p_{a_i} q_{x_i} \right) \left(\sqrt{p_n|_{a,\vec{a}_{k-1}} p_n|_{b,\vec{a}_{k-1}}} \right) \times$$
$$\left(\prod_{i=k+1}^{N} \left\langle \psi^{x_i,r_i}_{\vec{a}_i(a_k=b)} \middle| \psi^{x_i,r_i}_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} \right) = q_\mathrm{Z}\sqrt{p_a p_b} \sum_{a_1^{k-1}} \left(\prod_{i=1}^{k-1} p_{a_i} \right) \sqrt{p_n|_{a,\vec{a}_{k-1}} p_n|_{b,\vec{a}_{k-1}}} \times$$
$$\left[\sum_{a_{k+1}^N} \left(\prod_{i=k+1}^{N} p_{a_i} \left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} \right) \right] \tag{42}$$

for $k = 2, \ldots, N-1$, where we have made use of the fact that $\langle \psi^{x_i,r_i}_{\vec{a}_i(a_k=b)} | \psi^{x_i,r_i}_{\vec{a}_i(a_k=a)} \rangle_{B_i C_i}$ is independent of x_i and r_i for all i —which is straightforward to show from Eq. (35)— in order to carry out the sums over $\underline{x_k}$ and r_1^N: $\sum_{\underline{x_k}} \sum_{r_1^N} \left(\prod_{i\neq k} q_{x_i} \right) = \sum_{r_1^N} \left\{ \sum_{\underline{x_k}} \left(\prod_{i\neq k} q_{x_i} \right) \right\} = 2^N$. Also for this reason, in the last equality we have renamed $\langle \psi^{x_i,r_i}_{\vec{a}_i(a_k=b)} | \psi^{x_i,r_i}_{\vec{a}_i(a_k=a)} \rangle_{B_i C_i}$ simply as $\langle \psi_{\vec{a}_i(a_k=b)} | \psi_{\vec{a}_i(a_k=a)} \rangle_{B_i C_i}$. As expected, particularizing $a = b$ in Eq. (42), we obtain $\| |\widetilde{\Psi}^{(k)}_{a,\mathrm{Z},n}\rangle \|^2 = q_\mathrm{Z} p_a \sum_{a_1^{k-1}} (\prod_{i=1}^{k-1} p_{a_i}) p_n|_{a,\vec{a}_{k-1}}$. Thus, for the normalized states, we have

$$\left\langle \Psi^{(k)}_{b,\mathrm{Z},n} \middle| \Psi^{(k)}_{a,\mathrm{Z},n} \right\rangle = \sum_{a_1^{k-1}} \sqrt{p^{(k)}(\vec{a}_{k-1}|n,a,\mathrm{Z}) p^{(k)}(\vec{a}_{k-1}|n,b,\mathrm{Z})} \times$$
$$\left[\sum_{a_{k+1}^N} \left(\prod_{i=k+1}^{N} p_{a_i} \left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} \right) \right] \tag{43}$$

for $k = 2, \ldots, N-1$, where we have introduced the obvious definition

$$p^{(k)}(\vec{a}_{k-1}|n,a,\mathrm{Z}) = \frac{\left(\prod_{i=1}^{k-1} p_{a_i} \right) p_n|_{a,\vec{a}_{k-1}}}{\sum_{a_1^{k-1}} \left(\prod_{i=1}^{k-1} p_{a_i} \right) p_n|_{a,\vec{a}_{k-1}}}. \tag{44}$$

Lastly, regarding the extreme rounds (which are excluded from Eq. (43)), explicit calculation shows that

$$\left\langle \Psi^{(1)}_{b,\mathrm{Z},n} \middle| \Psi^{(1)}_{a,\mathrm{Z},n} \right\rangle = \sum_{a_2^N} \left(\prod_{i=2}^{N} p_{a_i} \left\langle \psi_{\vec{a}_i(a_1=b)} \middle| \psi_{\vec{a}_i(a_1=a)} \right\rangle_{B_i C_i} \right),$$
$$\left\langle \Psi^{(N)}_{b,\mathrm{Z},n} \middle| \Psi^{(N)}_{a,\mathrm{Z},n} \right\rangle = \sum_{a_1^{N-1}} \sqrt{p^{(N)}(\vec{a}_{N-1}|n,a,\mathrm{Z}) p^{(N)}(\vec{a}_{N-1}|n,b,\mathrm{Z})}. \tag{45}$$

We remark that, so far, we have not imposed Assumption 3 on the intensity correlations yet (see Sec. 2.1). At this stage, we invoke it by considering a finite correlation range ξ, which allows to rewrite Eq. (43) as

$$\left\langle \Psi^{(k)}_{b,\mathrm{Z},n} \middle| \Psi^{(k)}_{a,\mathrm{Z},n} \right\rangle =$$
$$\sum_{a_{\max\{k-\xi,1\}}^{k-1}} \sqrt{p^{(k)}(a_{k-1},\ldots,a_{\max\{k-\xi,1\}}|n,a,\mathrm{Z}) p^{(k)}(a_{k-1},\ldots,a_{\max\{k-\xi,1\}}|n,b,\mathrm{Z})}$$
$$\times \left[\sum_{a_{k+1}^{\min\{k+\xi,N\}}} \left(\prod_{i=k+1}^{\min\{k+\xi,N\}} p_{a_i} \left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} \right) \right] \tag{46}$$

for $k = 2, \ldots, N-1$, and similarly for Eq. (45). Crucially, we have made use of the fact that $\langle \psi_{\vec{a}_i(a_k=b)} | \psi_{\vec{a}_i(a_k=a)} \rangle_{B_i C_i} = 1$ for all $i > k+\xi$, which is a straightforward consequence of Assumption 3.

Aiming to evaluate Eq. (40), one must lower-bound the right-hand side of Eq. (46). For this purpose, we are going to exploit the structure of Eq. (3) in order to derive model-independent bounds valid for all correlations models $g_{\vec{a}_k}(\alpha_k)$. Let us address the bracket in Eq. (46) first. Noticing that $e^{-x}x^n$ is strictly decreasing (increasing) for $n = 0$ ($n = 1, 2, \ldots$) in the interval $x \in (0, 1)$, from Eq. (3) we have that $p_0|_{\vec{a}_i(a_k=a)} \geq e^{-a_i^+}$ and $p_{n\geq 1}|_{\vec{a}_i(a_k=a)} \geq e^{-a_i^-} a_i^{-\,n}/n!$ for all a, such that explicit calculation yields $\langle \psi_{\vec{a}_i(a_k=b)} | \psi_{\vec{a}_i(a_k=a)} \rangle_{B_i C_i} \geq 1 - (e^{-a_i^-} - e^{-a_i^+})$. Therefore, it is easy to show that

$$\sum_{a_{k+1}^{\min\{k+\xi,N\}}} \left(\prod_{i=k+1}^{\min\{k+\xi,N\}} p_{a_i} \left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} \right) \geq \left[1 - \sum_{c \in A} p_c \left(e^{-c^-} - e^{-c^+} \right) \right]^{\xi}, \tag{47}$$

which becomes a global prefactor in Eq. (46), as it does not depend on the remaining summation indexes $a_{\max\{k-\xi,1\}}$ to a_{k-1}. If we now focus on the first row of Eq. (46), the same monotonicity argument yields

$$p^{(k)}(a_{k-1}, \ldots, a_{\max\{k-\xi,1\}} | n, a, \mathrm{Z}) \geq \begin{cases} p_{a_{k-1}} \cdots p_{a_{\max\{k-\xi,1\}}} e^{a^- - a^+} & \text{if} \quad n = 0 \\ p_{a_{k-1}} \cdots p_{a_{\max\{k-\xi,1\}}} e^{a^+ - a^-} (a^-/a^+)^n & \text{if} \quad n \geq 1, \end{cases} \tag{48}$$

such that

$$\sum_{a_{\max\{k-\xi,1\}}^{k-1}} \sqrt{p^{(k)}(a_{k-1}, \ldots, a_{\max\{k-\xi,1\}} | n, a, \mathrm{Z}) p^{(k)}(a_{k-1}, \ldots, a_{\max\{k-\xi,1\}} | n, b, \mathrm{Z})} \geq \begin{cases} \exp\left\{ \frac{a^- + b^- - (a^+ + b^+)}{2} \right\} & \text{if} \quad n = 0 \\ \exp\left\{ \frac{a^+ + b^+ - (a^- + b^-)}{2} \right\} \left(\frac{a^- b^-}{a^+ b^+} \right)^{n/2} & \text{if} \quad n \geq 1 \end{cases} \tag{49}$$

for $k = 2, \ldots, N-1$. Now, putting together Eq. (47) and Eq. (49), we conclude that $| \langle \Psi_{b,\mathrm{Z},n}^{(k)} | \Psi_{a,\mathrm{Z},n}^{(k)} \rangle |^2 \geq \tau_{ab,n}^{\xi}$ for all $k = 2, \ldots N-1$, $n \in \mathbb{N}$, $a \in A$, $b \in A$ and $b \neq a$, where $\tau_{ab,n}^{\xi}$ is given in Eq. (6) of Sec. (2.2). Indeed, in virtue of Eq. (45), it is clear that the resulting bound also applies to the extreme rounds $k = 1$ and $k = N$ (the bound is simply less tight in these cases).

Lastly, to conclude the proof of Eq. (4), we need to establish the same result for the n-photon error click probabilities too, *i.e.,* we need to show that

$$G_- \left(H_{n,a,r}^{(k)}, \tau_{ab,n}^{\xi} \right) \leq H_{n,b,r}^{(k)} \leq G_+ \left(H_{n,a,r}^{(k)}, \tau_{ab,n}^{\xi} \right) \tag{50}$$

for all $k = 1, \ldots N$, $n \in \mathbb{N}$, $a \in A$, $b \in A$ and $b \neq a$, where we recall that $H_{n,a,r}^{(k)} = p^{(k)}(\mathrm{err} | n, a, \mathrm{X}, \mathrm{X}, r)$. For this purpose, note that, following identical steps as those leading to Eq. (39), one finds

$$H_{n,a,r}^{(k)} = \mathrm{Tr} \left\{ \hat{O}_{\mathrm{X},\mathrm{err},r}^{(k)} \left| \Psi_{a,\mathrm{X},r,n}^{(k)} \right\rangle \left\langle \Psi_{a,\mathrm{X},r,n}^{(k)} \right| \right\} = \left\langle \Psi_{a,\mathrm{X},r,n}^{(k)} \right| \hat{O}_{\mathrm{X},\mathrm{err},r}^{(k)} \left| \Psi_{a,\mathrm{X},r,n}^{(k)} \right\rangle \tag{51}$$

for

$$\hat{O}_{\mathrm{X},\mathrm{err},r}^{(k)} = \hat{U}_{BE}^{\dagger} \, \hat{M}_{B_k}^{\mathrm{X},1-r} \, \hat{U}_{BE} \quad \text{and} \quad \left| \Psi_{a,\mathrm{X},r,n}^{(k)} \right\rangle = \frac{\left| \widetilde{\Psi}_{a,\mathrm{X},r,n}^{(k)} \right\rangle}{\left\| \left| \widetilde{\Psi}_{a,\mathrm{X},r,n}^{(k)} \right\rangle \right\|}, \tag{52}$$

where $|\widetilde{\Psi}^{(k)}_{a,\mathrm{X},r,n}\rangle = \langle a,\mathrm{X}|_{A_k}\,\langle r|_{A'_k}\,\langle t_n|_{C_k}\,|\Psi\rangle$. Thus, in virtue of the CS constraint —given in Eq. (33)— we have that

$$G_-\left(H^{(k)}_{n,a,r},\left|\left\langle\Psi^{(k)}_{b,\mathrm{X},r,n}\middle|\Psi^{(k)}_{a,\mathrm{X},r,n}\right\rangle\right|^2\right) \leq H^{(k)}_{n,b,r} \leq G_+\left(H^{(k)}_{n,a,r},\left|\left\langle\Psi^{(k)}_{b,\mathrm{X},r,n}\middle|\Psi^{(k)}_{a,\mathrm{X},r,n}\right\rangle\right|^2\right), \quad (53)$$

and Eq. (50) follows from the fact that $|\langle\Psi^{(k)}_{b,\mathrm{X},r,n}|\Psi^{(k)}_{a,\mathrm{X},r,n}\rangle| = |\langle\Psi^{(k)}_{b,\mathrm{Z},n}|\Psi^{(k)}_{a,\mathrm{Z},n}\rangle|$ for both $r=0$ and $r=1$ and for any given n, k, a and b, which is easy to show following the same steps that lead to Eq. (43).

4.2 Phase error rate and secret key length in the finite key regime

The phase error rate is defined as $\phi_{1,\mathrm{Z},N} := E^{\mathrm{Z,ph}}_{1,\mu,N}/Z_{1,\mu,N}$, where $E^{\mathrm{Z,ph}}_{1,\mu,N}$ is the number of phase errors among all $Z_{1,\mu,N}$ single-photon events contributing to the sifted key. In this regard, we recall that a phase error is a bit error in a virtual entanglement-based protocol where, for the sifted key rounds, the parties measure their ancillas in the X basis instead. Let us define the set of rounds

$$\mathcal{M}_{1,\mu,N} = \mathcal{Z}_{1,\mu,N} \cup \mathcal{X}_{1,\mu,N}, \text{ where } \mathcal{Z}_{1,\mu,N} = \left\{k\middle|Z^{(k)}_{1,\mu}=1\right\} \text{ and } \mathcal{X}_{1,\mu,N} = \left\{k\middle|X^{(k)}_{1,\mu}=1\right\}. \quad (54)$$

The partition of $\mathcal{M}_{1,\mu,N}$ into $\mathcal{Z}_{1,\mu,N}$ and $\mathcal{X}_{1,\mu,N}$ is common to both the actual and the virtual protocol. In the virtual protocol, a specific number of bit errors occurs in $\mathcal{M}_{1,\mu,N}$, given by the number $E^{\mathrm{Z,ph}}_{1,\mu,N}$ of errors in $\mathcal{Z}_{1,\mu,N}$ plus the number $E_{1,\mu,N}$ of errors in $\mathcal{X}_{1,\mu,N}$. Basis-independence of the single-photon states delivered by Alice in the rounds indexed by $\mathcal{M}_{1,\mu,N}$ implies that Eve cannot distinguish test single-photons ($i \in \mathcal{X}_{1,\mu,N}$) from key single-photons ($i \in \mathcal{Z}_{1,\mu,N}$). Moreover, since, in the virtual protocol, Alice and Bob measure their ancillas in the X basis in both types of rounds, for any given round in $\mathcal{M}_{1,\mu,N}$ the probability that it yields an error is independent of its round-type in the virtual protocol. Thus, one can imagine that Eve is inducing the bit errors in $\mathcal{M}_{1,\mu,N}$ first, and later on Alice and Bob randomly select a partition $\mathcal{M}_{1,\mu,N} = \mathcal{Z}_{1,\mu,N} \cup \mathcal{X}_{1,\mu,N}$, such that

$$\langle\phi_{1,\mathrm{Z},N}\rangle = \left\langle\frac{E_{1,\mu,N}}{X_{1,\mu,N}}\right\rangle \quad (55)$$

and one can derive a statistical upper bound on the difference $|\phi_{1,\mathrm{Z},N} - E_{1,\mu,N}/X_{1,\mu,N}|$ via Serfling's inequality [31]. For our purposes, the relevant one-sided bound can be stated as [30]

$$P\left\{\phi_{1,\mathrm{Z},N} > \frac{\overline{E}_{1,\mu,N}}{\overline{X}_{1,\mu,N}} + \gamma_{\epsilon_\mathrm{S}}\right\} \leq \epsilon_\mathrm{S} \text{ for } \gamma_{\epsilon_\mathrm{S}} = \sqrt{\frac{\left(\overline{X}_{1,\mu,N}+\overline{Z}_{1,\mu,N}\right)\left(\overline{Z}_{1,\mu,N}+1/N\right)}{2N\overline{Z}^2_{1,\mu,N}\overline{X}_{1,\mu,N}}\log\left(\frac{1}{\epsilon_\mathrm{S}}\right)}. \quad (56)$$

Now, let the following inequalities be given:

$$P(\overline{Z}_{1,\mu,N} < \overline{Z}^{\mathrm{L},\epsilon_1}_{1,\mu,N}) \leq \epsilon_1, \;\; P(\overline{X}_{1,\mu,N} < \overline{X}^{\mathrm{L},\epsilon_2}_{1,\mu,N}) \leq \epsilon_2, \text{ and } P(\overline{E}_{1,\mu,N} > \overline{E}^{\mathrm{U},\epsilon_3}_{1,\mu,N}) \leq \epsilon_3, \quad (57)$$

for certain $\overline{Z}^{\mathrm{L},\epsilon_1}_{1,\mu,N}$, $\overline{X}^{\mathrm{L},\epsilon_2}_{1,\mu,N}$ and $\overline{E}^{\mathrm{U},\epsilon_3}_{1,\mu,N}$ that depend on the observables. In virtue of the union bound, it follows from Eq. (56) and Eq. (57) that

$$P\left\{\phi_{1,\mathrm{Z},N} > \frac{E^{\mathrm{U},\epsilon_3}_{1,\mu,N}}{X^{\mathrm{L},\epsilon_2}_{1,\mu,N}} + \gamma_{\epsilon_\mathrm{S},\epsilon_1,\epsilon_2}\right\} \leq \epsilon_\mathrm{S} + \sum_{i=1}^{3}\epsilon_i \quad (58)$$

for

$$\gamma_{\epsilon_{\mathrm{S}},\epsilon_1,\epsilon_2} = \sqrt{\frac{\left(\overline{X}^{\mathrm{L},\epsilon_2}_{1,\mu,N} + \overline{Z}^{\mathrm{L},\epsilon_1}_{1,\mu,N}\right)\left(\overline{Z}^{\mathrm{L},\epsilon_1}_{1,\mu,N} + 1/N\right)}{2N\overline{Z}^{\mathrm{L},\epsilon_1\,2}_{1,\mu,N}\overline{X}^{\mathrm{L},\epsilon_2}_{1,\mu,N}}\log\left(\frac{1}{\epsilon_{\mathrm{S}}}\right)}. \tag{59}$$

At this stage, one can present the secret key rate of the decoy-state BB84 protocol under consideration, which relies on a lower bound on $\overline{Z}_{1,\mu,N}$ (presumed in Eq. (57)) and an upper bound on $\phi_{1,\mathrm{Z},N}$ (Eq. (58)). Precisely, for any $\delta \in (0,1)$, it is known that privacy amplification allows to extract an ϵ_{sec}-secret, ϵ_{cor}-correct secret key of length [15]

$$l = Z^{\mathrm{L},\epsilon_1}_{1,\mu,N}\left[1 - h\left(\frac{E^{\mathrm{U},\epsilon_3}_{1,\mu,N}}{X^{\mathrm{L},\epsilon_2}_{1,\mu,N}} + \gamma_{\epsilon_{\mathrm{S}},\epsilon_1,\epsilon_2}\right)\right] - f_{\mathrm{EC}}Z_{\mu,N}h(E_{\mathrm{tol}}) - \log\left(\frac{1}{\epsilon_{\mathrm{cor}}\epsilon^2_{\mathrm{PA}}\delta}\right) \tag{60}$$

as long as $\epsilon_{\mathrm{sec}} \geq 2\varepsilon + \epsilon_{\mathrm{PA}} + \delta$, where f_{EC} is the efficiency of the error correction protocol, $Z_{\mu,N}$ provides the length of the sifted key (see Sec. 2.3 for the definition of $Z_{\mu,N}$), $h(\cdot)$ denotes the binary entropy function, E_{tol} is a threshold bit error rate for the error correction, ϵ_{PA} is the error probability of the privacy amplification and $\varepsilon = \epsilon_{\mathrm{S}} + \sum_{i=1}^{3}\epsilon_i$ is the parameter estimation error, *i.e.*, an upper bound on the total error probability of the parameter estimation. Of course, the secret key rate is defined as $K_N = l/N$. That is to say,

$$K_N = \overline{Z}^{\mathrm{L},\epsilon_1}_{1,\mu,N}\left[1 - h\left(\frac{\overline{E}^{\mathrm{U},\epsilon_3}_{1,\mu,N}}{\overline{X}^{\mathrm{L},\epsilon_2}_{1,\mu,N}} + \gamma_{\epsilon_{\mathrm{S}},\epsilon_1,\epsilon_2}\right)\right] - f_{\mathrm{EC}}\overline{Z}_{\mu,N}h(E_{\mathrm{tol}}) - \frac{1}{N}\log\left(\frac{1}{\epsilon_{\mathrm{cor}}\epsilon^2_{\mathrm{PA}}\delta}\right), \tag{61}$$

where we have introduced the notation $\overline{Y} = Y/N$.

4.3 Technical claims on the asymptotic regime

The asymptotic secret key rate formula given in Sec. 2.5 builds on the assertion that, as long as the variance of the experimental averages tends to zero as $N \to \infty$, the probability of any finite violation of Eq. (28) vanishes for $N \to \infty$ too. Propositions 1 and 2 below formally demonstrate this claim.

Proposition 1. Let us assume that $\lim_{N\to\infty} Var\left[\overline{Z}_{a,N}\right] = 0$ for all $a \in A$ and $\lim_{N\to\infty} Var\left[\overline{Z}_{1,\mu,N}\right] = 0$. Then, $\lim_{N\to\infty} P(\overline{Z}_{1,\mu,N} \leq \overline{Z}^{\mathrm{L}}_{1,\mu,N} - \delta) = 0$ for all $\delta > 0$. The proposition holds too if one replaces Z by X everywhere.

Proof. Let us consider the event $E_{\delta,N} = \left\{\overline{Z}^{\mathrm{L}}_{1,\mu,N} - \overline{Z}_{1,\mu,N} \geq \delta\right\}$. We have

$$\begin{aligned}
E_{\delta,N} &= \left\{\overline{Z}^{\mathrm{L}}_{1,\mu,N} - \langle\overline{Z}^{\mathrm{L}}_{1,\mu,N}\rangle + \left\langle\overline{Z}_{1,\mu,N}\right\rangle - \overline{Z}_{1,\mu,N} + \langle\overline{Z}^{\mathrm{L}}_{1,\mu,N}\rangle - \left\langle\overline{Z}_{1,\mu,N}\right\rangle \geq \delta\right\} \\
&\subseteq \left\{\left|\overline{Z}^{\mathrm{L}}_{1,\mu,N} - \langle\overline{Z}^{\mathrm{L}}_{1,\mu,N}\rangle\right| + \left|\left\langle\overline{Z}_{1,\mu,N}\right\rangle - \overline{Z}_{1,\mu,N}\right| + \langle\overline{Z}^{\mathrm{L}}_{1,\mu,N}\rangle - \left\langle\overline{Z}_{1,\mu,N}\right\rangle \geq \delta\right\} \\
&\subseteq \left\{\left|\overline{Z}^{\mathrm{L}}_{1,\mu,N} - \langle\overline{Z}^{\mathrm{L}}_{1,\mu,N}\rangle\right| + \left|\left\langle\overline{Z}_{1,\mu,N}\right\rangle - \overline{Z}_{1,\mu,N}\right| \geq \delta\right\} \\
&\subseteq \left\{\left|\overline{Z}^{\mathrm{L}}_{1,\mu,N} - \langle\overline{Z}^{\mathrm{L}}_{1,\mu,N}\rangle\right| \geq \frac{\delta}{2}\right\} \cup \left\{\left|\left\langle\overline{Z}_{1,\mu,N}\right\rangle - \overline{Z}_{1,\mu,N}\right| \geq \frac{\delta}{2}\right\}
\end{aligned} \tag{62}$$

where in the first set bound we used the triangle inequality twice, in the second one we used the fact that $\langle\overline{Z}^{\mathrm{L}}_{1,\mu,N}\rangle - \left\langle\overline{Z}_{1,\mu,N}\right\rangle \leq 0$ for all N —according to the first decoy-state

bound in Eq. (26)— and in the third one we used the fact that, if $|X| + |Y| \geq \delta$, then either $|X| \geq \delta/2$ or $|Y| \geq \delta/2$. Now, in virtue of the union bound, we have that

$$P\left(E_{\delta,N}\right) \leq P\left(|\overline{Z}^{\mathrm{L}}_{1,\mu,N} - \langle \overline{Z}^{\mathrm{L}}_{1,\mu,N} \rangle| \geq \frac{\delta}{2}\right) + P\left(\left|\left\langle \overline{Z}_{1,\mu,N} \right\rangle - \overline{Z}_{1,\mu,N}\right| \geq \frac{\delta}{2}\right). \tag{63}$$

Therefore, the claim holds if we show that both terms in the right-hand side tend to zero as N tends to infinity for all $\delta > 0$. Recalling that, in virtue of Chebyshev's inequality [32], mean-square convergence of a sequence of random variables guarantees convergence in probability, for the second term of Eq. (63) we have

$$\lim_{N\to\infty} Var\left[\overline{Z}_{1,\mu,N}\right] = 0 \implies \lim_{N\to\infty} P\left(\left|\overline{Z}_{1,\mu,N} - \left\langle \overline{Z}_{1,\mu,N} \right\rangle\right| \geq \frac{\delta}{2}\right) = 0 \tag{64}$$

for all $\delta > 0$. Regarding the first term, note that $\overline{Z}^{\mathrm{L}}_{1,\mu,N}$ is linear in the $\overline{Z}_{a,N}$ (see Sec. 2.4). That is to say, $\overline{Z}^{\mathrm{L}}_{1,\mu,N} = \sum_{a\in A} c_a \overline{Z}_{a,N} + C$ for certain coefficients c_a and C. Thus,

$$Var\left[\overline{Z}^{\mathrm{L}}_{1,\mu,N}\right] = E\left[\left|\sum_{a\in A} c_a \left(\overline{Z}_{a,N} - \langle \overline{Z}_{a,N} \rangle\right)\right|^2\right], \tag{65}$$

and since [32] $E[|X+Y|^2] \leq 4(E[|X|^2] + E[|Y|^2])$, it follows that

$$Var\left[\overline{Z}^{\mathrm{L}}_{1,\mu,N}\right] \leq K \sum_a |c_a|^2 Var\left[\overline{Z}_{a,N}\right] \tag{66}$$

for some positive constant K, such that

$$\begin{aligned} &\lim_{N\to\infty} Var\left[\overline{Z}_{a,N}\right] = 0 \text{ for all } a \in A \implies \lim_{N\to\infty} Var\left[\overline{Z}^{\mathrm{L}}_{1,\mu,N}\right] = 0 \implies \\ &\lim_{N\to\infty} P\left(|\overline{Z}^{\mathrm{L}}_{1,\mu,N} - \langle \overline{Z}^{\mathrm{L}}_{1,\mu,N} \rangle| \geq \frac{\delta}{2}\right) = 0 \end{aligned} \tag{67}$$

for all $\delta > 0$, where again we invoked the fact that mean-square convergence implies convergence in probability. □

Proposition 2. Let us assume that $\lim_{N\to\infty} Var\left[\overline{E}_{a,N}\right] = 0$ for all $a \in A$ and $\lim_{N\to\infty} Var\left[\overline{E}_{1,\mu,N}\right] = 0$. Then, $\lim_{N\to\infty} P(\overline{E}_{1,\mu,N} \geq \overline{E}^{\mathrm{U}}_{1,\mu,N} + \delta) = 0$ for all $\delta > 0$.

The proof of Proposition 2 follows identically as that of Proposition 1.

As a final comment, note that, when dealing with bounded sequences of random variables —$\{X_j\}$ is bounded if there exists some constant C such that $|X_j| < C$ for all j— mean-square convergence is not stronger but exactly equivalent to convergence in probability (see for instance [32]), such that demanding the latter kind of convergence instead does not relax the preconditions of propositions 1 and 2. If, alternatively, neither kind of convergence is demanded, all we know is that Eq. (26) holds for the expectations, which does not suffice to establish the limits of propositions 1 and 2.

5 Data availability

No datasets were generated or analysed during the current study.

6 Acknowledgements

We thank Margarida Pereira for very fruitful discussions. This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675662 (project QCALL), by the Galician Regional Government (consolidation of Research Units: AtlantTIC), the Spanish Ministry of Economy and Competitiveness (MINECO), the Fondo Europeo de Desarrollo Regional (FEDER) through Grant No. PID2020-118178RB-C21, and the Spanish Ministry of Science and Innovation through the "Planes Complementarios de I+D+I con las Comunidades Autónomas" in Quantum Communication. V.Z. and A.N. acknowledge support from respective FPU pre-doctoral scholarships from the Spanish Ministry of Education. K.T. acknowledges support from JSPS KAKENHI Grant Numbers JP18H05237 18H05237 and JST-CREST JPMJCR 1671.

7 Author contributions

M.C. and K.T. conceived the initial idea and triggered the consideration of this project. V.Z. and A.N. made the theoretical analysis and performed the numerical simulations, with inputs from all authors. All authors analysed the results and prepared the manuscript.

8 Competing interests

The authors declare no competing interests.

References

[1] Scarani, V. *et al.* The security of practical quantum key distribution. *Rev. Mod. Phys.* **81,** 1301 (2009).

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

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

[4] Vernam, G. S., *Trans. Am. Inst. Electr. Eng.* XLV, 295 (1926).

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

[6] Yin, H. L., *et al.* Measurement-device-independent quantum key distribution over a 404 km optical fiber. *Phys. Rev. Lett.* **117,** 190501 (2016).

[7] Boaron, A., *et al.* Secure quantum key distribution over 421 km of optical fiber. *Phys. Rev. Lett.* **121,** 190502 (2018).

[8] Fang, X. T., *et al.* Implementation of quantum key distribution surpassing the linear rate-transmittance bound. *Nat. Photonics* **14,** 422-425 (2020).

[9] Chen, J. P., *et al.* Sending-or-not-sending with independent lasers: secure twin-field quantum key distribution over 509 km. *Phys. Rev. Lett.* **124,** 070501 (2020).

[10] Yoshino, K. I. *et al.* Quantum key distribution with an efficient countermeasure against correlated intensity fluctuations in optical pulses. *npj Quantum Inf.* **4,** 1-8 (2018).

[11] Grünenfelder, F., Boaron, A., Rusca, D., Martin, A. & Zbinden, H. Performance and security of 5 GHz repetition rate polarization-based quantum key distribution. *Appl. Phys. Lett.* **117,** 144003 (2020).

[12] Hwang, W. Y. Quantum key distribution with high loss: toward global secure communication. *Phys. Rev. Lett.* **91,** 057901 (2003).

[13] Lo, H.-K., Ma, X. & Chen, K. Decoy state quantum key distribution. *Phys. Rev. Lett.* **94,** 230504 (2005).

[14] Wang, X.-B. Beating the photon-number-splitting attack in practical quantum cryptography. *Phys. Rev. Lett.* **94,** 230503 (2005).

[15] Lim, C. C. W., Curty, M., Walenta, N., Xu, F. & Zbinden, H. Concise security bounds for practical decoy-state quantum key distribution. *Phys. Rev. A* **89,** 022307 (2014).

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

[17] Nagamatsu Y., Mizutani, A., Ikuta, R., Yamamoto, T., Imoto, N., & Tamaki, K. Security of quantum key distribution with light sources that are not independently and identically distributed. *Phys. Rev. A* **93,** 042325 (2016).

[18] Mizutani, A. *et al.* Quantum key distribution with setting-choice-independently correlated light sources. *npj Quantum Inf.* **5,** 8 (2019).

[19] Roberts, G. L. *et al.* Patterning-effect mitigating intensity modulator for secure decoy-state quantum key distribution. *Optics letters* **43,** 5110-5113 (2018).

[20] Pereira, M., Kato, G., Mizutani, A., Curty, M. & Tamaki, K. Quantum key distribution with correlated sources. *Science Advances* **6,** eaaz4487 (2020).

[21] Lo, H.-K., Curty, M., & Qi, B. Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.* **108,** 130503 (2012).

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

[23] Lo, H.-K. & Preskill, J. Security of quantum key distribution using weak coherent states with nonrandom phases. *Quantum Inf. Comput.* **7,** 431–458 (2007).

[24] Mitzenmacher, M., & Upfal, E. Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis (Cambridge University Press, 2017).

[25] Hoeffding, W. Probability inequalities for sums of bounded random variables. *J. Am. Stat. Assoc.* **58,** 13-30 (1963).

[26] Zapatero, V., & Curty, M. Secure quantum key distribution with a subset of malicious devices. *npj Quantum Inf.* 7, 1-8 (2021).

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

[28] Navarrete, Á., Pereira, M., Curty, M., & Tamaki K. Practical quantum key distribution that is secure against side channels. *Phys. Rev. Appl.* **15,** 034072 (2021).

[29] Bazaraa, M. S., Jarvis, J. J., & Sherali, H. D. *Linear programming and network flows*, John Wiley & Sons (2008).

[30] Tomamichel, M., Lim, C. C. W., Gisin, N., & Renner, R. Tight finite-key analysis for quantum cryptography. *Nat. Commun.* **3,** 1-6 (2012).

[31] Serfling, R. J. Probability inequalities for the sum in sampling without replacement. *Ann. Stat.*, 39-48 (1974).

[32] Billingsley, P. *Convergence of probability measures*, John Wiley & Sons (2013).

A Reference values for the linearized Cauchy-Schwarz constraints

Below, we provide the reference values $\tilde{y}_{n,a}$ and $\tilde{h}_{n,a}$ that follow from the typical channel model presented in Sec. 2.6 of the main text, which depends on the experimental inputs η, δ_{A} and p_{d}.

For convenience, we calculate $\tilde{h}_{n,a}$ first, for which we proceed in two steps. Disregarding the dark counts and the random assignments of the double clicks for the moment, the possible genuine detection outcomes for an n-photon pulse emitted by Alice are "no click", "no error", "error" and "double click", respectively denoted as 00, 10, 01 and 11. Their probabilities are

$$\begin{aligned} p_{00} &= (1-\eta)^n, \\ p_{10} &= \left(\eta \cos^2 \delta_{\mathrm{A}} + 1 - \eta\right)^n - (1-\eta)^n, \\ p_{01} &= \left(\eta \sin^2 \delta_{\mathrm{A}} + 1 - \eta\right)^n - (1-\eta)^n, \\ p_{11} &= 1 - p_{00} - p_{01} - p_{10}. \end{aligned} \tag{68}$$

Eq. (68) can be interpreted as follows: every photon in the n-photon Fock state emitted by Alice's PRWCPs source reaches Bob's lab with probability η and experiences a polarization bit flip with probability $\sin^2 \delta_{\mathrm{A}}$, as illustrated in Fig. 2.

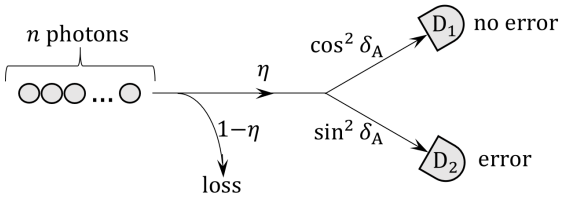

Figure 2: Schematic representation of the typical channel model considered in Sec. 2.6 of the main text. We recall that n stands for the number of photons emitted by Alice's PRWCP source, η stands for the overall system efficiency and δ_{A} stands for the polarization misalignment.

In order to incorporate the dark counts and the random assignments of the double clicks, we introduce the mutually exclusive events $A = \{\text{no dark counts}\}$, $B = \{\text{dark count in } D_1\}$, $C = \{\text{dark count in } D_2\}$ and $D = \{\text{dark count in both } D_1 \text{ and } D_2\}$, where we follow the detector notation of Fig. 2. The conditional error probabilities read

$$\begin{aligned} p_{\mathrm{err}}|_A &= p_{01} + \frac{1}{2} p_{11}, \\ p_{\mathrm{err}}|_B &= \frac{1}{2}\left(p_{01} + p_{11}\right), \\ p_{\mathrm{err}}|_C &= p_{00} + p_{01} + \frac{1}{2}\left(p_{10} + p_{11}\right), \\ p_{\mathrm{err}}|_D &= \frac{1}{2}, \end{aligned} \tag{69}$$

and, consequently,

$$\tilde{h}_{n,a} = (1-p_{\mathrm{d}})^2 p_{\mathrm{err}}|_A + p_{\mathrm{d}}(1-p_{\mathrm{d}})\left(p_{\mathrm{err}}|_B + p_{\mathrm{err}}|_C\right) + p_{\mathrm{d}}^2 p_{\mathrm{err}}|_D \tag{70}$$

for all $n \in \mathbb{N}$ and $a \in A$. Of course, regarding $\tilde{y}_{n,a}$, we have

$$\tilde{y}_{n,a} = 1 - (1-p_{\mathrm{d}})^2 p_{00} \tag{71}$$

for all $n \in \mathbb{N}$ and $a \in A$.

B Trace distance argument

The trace distance (TD) argument is stated as follows.

Theorem [27]. Let ρ and σ be two distinct states of a certain quantum system. Then, the trace distance between ρ and σ satisfies $D(\rho,\sigma) = \max\{\mathrm{Tr}(\hat{O}(\rho-\sigma))\}$, where the maximization is taken over all positive operators $\hat{O} \leq I$.

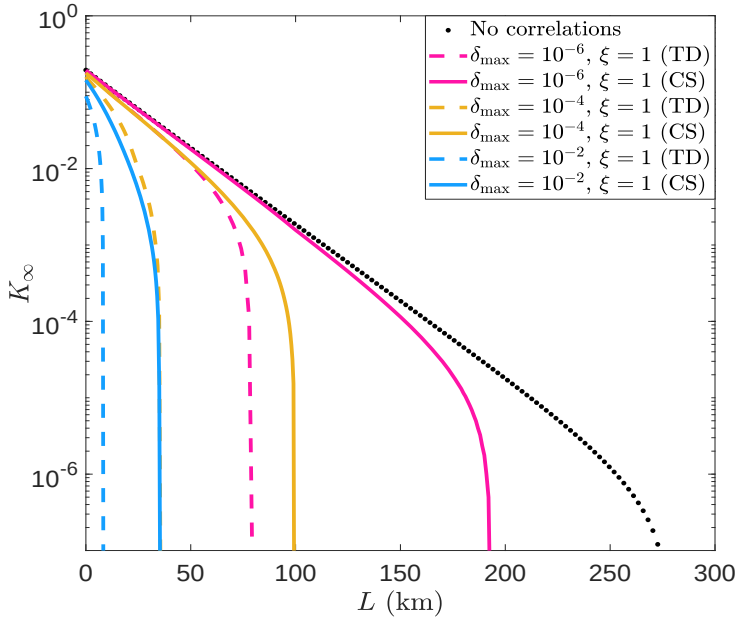

Figure 3: Comparison between the TD argument and the linearized CS constraint in terms of their secret key rate performance. For illustration purposes, only nearest-neighbors intensity correlations are contemplated, *i.e.*, we set $\xi = 1$. On the one hand, the dashed lines are obtained using the TD argument, showing the asymptotic secret key rate, K_∞, as a function of the distance, L, for various values of the maximum relative deviation between intensity settings (a_k) and actual intensities (α_k), $\delta_{\max} \in \{10^{-6}, 10^{-4}, 10^{-2}\}$. On the other hand, the solid lines represent the corresponding secret key rates obtained with the CS inequality instead. Although these latter lines also appear in Fig. 1 of the main text, for clarity purposes the color and line style criteria are different here. In addition, we include the attainable secret key rate in the absence of intensity correlations for completeness (dotted black line). Regarding the experimental parameters, they are fixed identically as in Fig. 1 of the main text.

Keeping the notation $Y^{(k)}_{n,a} = p^{(k)}(\mathrm{click}|n,a,\mathrm{Z},\mathrm{Z})$, and making use of the fact that $D(|x\rangle\langle x|, |y\rangle\langle y|) = (1-|\langle x|y\rangle|^2)^{1/2}$ [27], the bound provided by the TD argument reads [16]

$$\left|Y^{(k)}_{n,a} - Y^{(k)}_{n,b}\right| \leq \sqrt{1-\left|\left\langle \Psi^{(k)}_{b,\mathrm{Z},n} \middle| \Psi^{(k)}_{a,\mathrm{Z},n}\right\rangle\right|^2} \leq \sqrt{1-\tau^{\xi}_{ab,n}} \tag{72}$$

for all $a \in A$, $b \in A$ $(b \neq a)$, $n \in \mathbb{N}$ and $k = 1, \ldots, N$. Here, we have used the lower bound

$$\left|\left\langle \Psi^{(k)}_{b,\mathrm{Z},n} \middle| \Psi^{(k)}_{a,\mathrm{Z},n}\right\rangle\right|^2 \geq \tau^{\xi}_{ab,n} \tag{73}$$

presented in the main text, which depends on a presumed finite correlation range ξ. Remarkably, Eq. (72) does not rely on a characterization of the quantum channel, as opposed to the linearized CS constraints. What is more, the TD argument provides equivalent constraints for the n-photon error click probabilities too, as seen next. In the first place,

$$\left|H^{(k)}_{n,a,r} - H^{(k)}_{n,b,r}\right| \leq \sqrt{1-\left|\left\langle \Psi^{(k)}_{b,\mathrm{X},r,n} \middle| \Psi^{(k)}_{a,\mathrm{X},r,n}\right\rangle\right|^2} \tag{74}$$

for $n \in \mathbb{N}$, $a \in A$, $b \in A$ $(b \neq a)$, $r \in \mathbb{Z}_2$ and $k = 1, \ldots, N$, where we maintain the notation $H^{(k)}_{n,a,r} = p^{(k)}(\mathrm{err}|n, a, \mathrm{X}, \mathrm{X}, r)$ and $H^{(k)}_{n,a} = p^{(k)}(\mathrm{err}|n, a, \mathrm{X}, \mathrm{X})$. If, in addition, we recall that $|\langle \Psi^{(k)}_{b,\mathrm{X},r,n} | \Psi^{(k)}_{a,\mathrm{X},r,n} \rangle| = |\langle \Psi^{(k)}_{b,\mathrm{Z},n} | \Psi^{(k)}_{a,\mathrm{Z},n} \rangle|$ for both $r = 0$ and $r = 1$, the desired bound follows from the triangle inequality:

$$\left| H^{(k)}_{n,a} - H^{(k)}_{n,b} \right| = \left| \frac{1}{2} \left(H^{(k)}_{n,a,0} + H^{(k)}_{n,a,1} \right) - \frac{1}{2} \left(H^{(k)}_{n,b,0} + H^{(k)}_{n,b,1} \right) \right| \leq$$
$$\frac{1}{2} \left| H^{(k)}_{n,a,0} - H^{(k)}_{n,b,0} \right| + \frac{1}{2} \left| H^{(k)}_{n,a,1} - H^{(k)}_{n,b,1} \right| \leq \sqrt{1 - \tau^{\xi}_{ab,n}} \tag{75}$$

for any given finite correlation range ξ.

Aiming to compare the TD argument and the linearized CS constraints in terms of their secret key rate performance, one must replace the corresponding restrictions by Eq. (72) and Eq. (75) in the linear programs of Sec. 2.4. The result is illustrated in Fig. 3 for the most representative case of nearest-neighbors intensity correlations, *i.e.*, $\xi = 1$.

Comparing Fig. 3 with Fig. 1 in the main text, we see that the linearized CS constraint provides significantly tighter bounds than the TD argument for the parameter estimation, as long as adequate reference parameters are given as inputs to the former.

C Deterministic intensity correlations model

In this note, we consider a deterministic intensity correlations model where, at every round k, the record of settings $(\vec{a}_k)$ fully determines the intensity (α_k), instead of just pinning its probability distribution. Nevertheless, we assume that the exact value of α_k is unknown to keep the analysis as general as possible. That is to say, for any given record $\vec{a}_k$, the model reads

$$g_{\vec{a}_k}(\alpha_k) = \delta \left(\alpha_k - \alpha_k^{\vec{a}_k} \right), \tag{76}$$

for some unknown $\alpha_k^{\vec{a}_k} \in [a_k^-, a_k^+]$ fixed by $\vec{a}_k$ (the worst case will be considered), where $\delta(\cdot)$ stands for the Dirac delta distribution. This model allows to compute a tighter lower bound for the overlap $|\langle \Psi^{(k)}_{b,\mathrm{Z},n} | \Psi^{(k)}_{a,\mathrm{Z},n} \rangle|$ than the one derived in the model-independent case. For this purpose, the starting point is the equation

$$\left\langle \Psi^{(k)}_{b,\mathrm{Z},n} \middle| \Psi^{(k)}_{a,\mathrm{Z},n} \right\rangle =$$
$$\sum_{a^{k-1}_{\max\{k-\xi,1\}}} \sqrt{p^{(k)}(a_{k-1}, \ldots, a_{\max\{k-\xi,1\}}|n, a, \mathrm{Z}) p^{(k)}(a_{k-1}, \ldots, a_{\max\{k-\xi,1\}}|n, b, \mathrm{Z})}$$
$$\times \left[\sum_{a^{\min\{k+\xi,N\}}_{k+1}} \left(\prod_{i=k+1}^{\min\{k+\xi,N\}} p_{a_i} \left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} \right) \right] \tag{77}$$

for $k = 2, \ldots, N-1$, where we recall that

$$\left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} = \sum_{n=0}^{\infty} \left(p_n|_{\vec{a}_i(a_k=b)} \times p_n|_{\vec{a}_i(a_k=a)} \right)^{1/2} \tag{78}$$

for all $i = k+1, \ldots, N$, and ξ stands for the finite correlation range. Noticing that Eq. (76) implies

$$p_n|_{\vec{a}_i(a_k=b)} = \frac{\exp\left\{ -\alpha_i^{\vec{a}_i(a_k=b)} \right\} \left(\alpha_i^{\vec{a}_i(a_k=b)} \right)^n}{n!} \tag{79}$$

for a fixed (but unknown) $\alpha_i^{\vec{a}_i(a_k=b)} \in [a_i^-, a_i^+]$, $a_i \in A$, it follows that

$$\begin{aligned}
&\left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} = \\
&\sum_{n=0}^{\infty} \exp\left\{ -\left(\alpha_i^{\vec{a}_i(a_k=a)} + \alpha_i^{\vec{a}_i(a_k=b)} \right)/2 \right\} \sqrt{\alpha_i^{\vec{a}_i(a_k=a)} \alpha_i^{\vec{a}_i(a_k=b)}}^{\,n} / n! = \\
&\exp\left\{ \sqrt{\alpha_i^{\vec{a}_i(a_k=a)} \alpha_i^{\vec{a}_i(a_k=b)}} - \left(\alpha_i^{\vec{a}_i(a_k=a)} + \alpha_i^{\vec{a}_i(a_k=b)} \right)/2 \right\}. \qquad (80)
\end{aligned}$$

Now, analytically minimizing this overlap for $\left(\alpha_i^{\vec{a}_i(a_k=a)}, \alpha_i^{\vec{a}_i(a_k=b)} \right) \in [a_i^-, a_i^+] \times [a_i^-, a_i^+]$, one obtains

$$\left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} \geq \exp\left\{ \sqrt{a_i^+ a_i^-} - (a_i^+ + a_i^-)/2 \right\}, \qquad (81)$$

such that the bracket in Eq. (77) is lower-bounded as

$$\begin{aligned}
&\sum_{a_{k+1}^{\min\{k+\xi,N\}}} \left(\prod_{i=k+1}^{\min\{k+\xi,N\}} p_{a_i} \left\langle \psi_{\vec{a}_i(a_k=b)} \middle| \psi_{\vec{a}_i(a_k=a)} \right\rangle_{B_i C_i} \right) \geq \\
&\left[\sum_{c \in A} p_c \exp\left\{ \sqrt{c^+ c^-} - (c^+ + c^-)/2 \right\} \right]^{\xi}, \qquad (82)
\end{aligned}$$

which factors off the remaining summations of Eq. (77). For the latter, we maintain the model-independent bound provided in the main text for simplicity. Thus, putting both terms together and recalling that the resulting bound also applies to the extreme rounds $k = 1$ and $k = N$ (this consequence follows identically as in the Methods Sec. 4.1), we conclude that

$$\begin{aligned}
&\left| \left\langle \Psi_{b,\mathrm{Z},n}^{(k)} \middle| \Psi_{a,\mathrm{Z},n}^{(k)} \right\rangle \right|^2 \geq \\
&\gamma_{ab,n}^{\xi} = \begin{cases} e^{a^- + b^- - (a^+ + b^+)} \left[\sum_{c \in A} p_c \exp\left\{ \sqrt{c^+ c^-} - (c^+ + c^-)/2 \right\} \right]^{2\xi} & \text{if} \quad n = 0 \\ e^{a^+ + b^+ - (a^- + b^-)} \left(\frac{a^- b^-}{a^+ b^+} \right)^n \left[\sum_{c \in A} p_c \exp\left\{ \sqrt{c^+ c^-} - (c^+ + c^-)/2 \right\} \right]^{2\xi} & \text{if} \quad n \geq 1 \end{cases} \qquad (83)
\end{aligned}$$

for all $k = 1, \ldots N$, $n \in \mathbb{N}$, $a \in A$, $b \in A$ and $b \neq a$.

Remarkably, the only difference introduced by the deterministic model in the parameter estimation procedure consists of replacing $\tau_{ab,n}^{\xi}$ by $\gamma_{ab,n}^{\xi}$ everywhere in the linearized CS constraints. Beyond this, the linear programs of Sec. 2.4 remain unchanged.

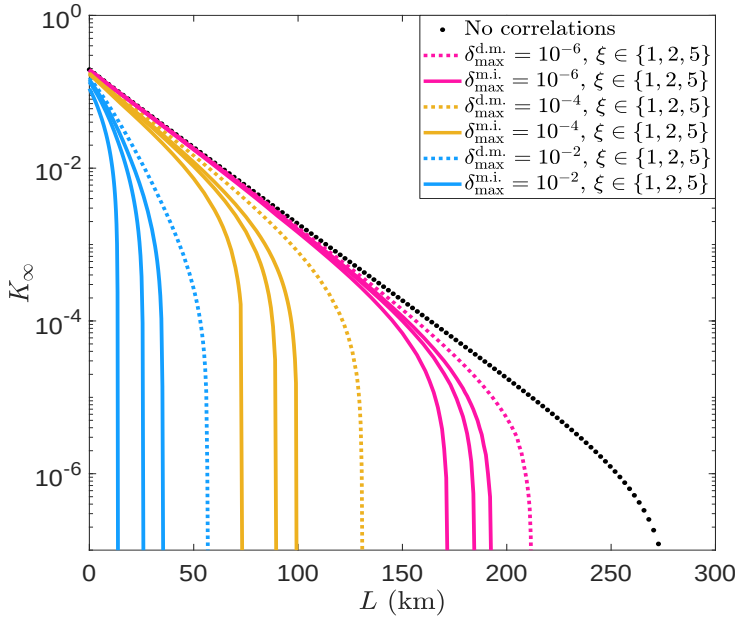

Figure 4: Comparison between deterministic intensity correlations (denoted by "d.m." in the figure legend) and model-independent intensity correlations (denoted by "m.i." in the figure legend) in terms of their secret key rate performance. The dotted color lines show the asymptotic secret key rate, K_∞, as a function of the distance, L, assuming the deterministic model. In accordance with Fig. 3, we contemplate three different values of the maximum relative deviation between intensity settings (a_k) and actual intensities (α_k), $\delta_{\max} \in \{10^{-6}, 10^{-4}, 10^{-2}\}$. Similarly, the solid color lines represent the corresponding secret key rates in the model-independent scenario. Importantly, three correlation ranges are used, namely, $\xi \in \{1,2,5\}$. However, the effect of this parameter on the secret key rate is negligible in the deterministic model for such moderate values, and thus we only plot $\xi = 1$ in that case. As a reference, we provide the attainable key rate in the absence of intensity correlations too (black dotted line). Regarding the experimental parameters, they are common with those of Fig. 1 in the main text.

To finish with, Fig. 4 illustrates how the asymptotic secret key rate K_∞ is enhanced when one moves from the model-independent scenario of Fig. 1 in the main text to a deterministic model. Remarkably, as seen in the figure, this model is also more robust to the correlation range ξ than the model-independent setting.