Hacking coherent-one-way quantum key distribution with present-day technology

To cite this article: Javier Rey-Domínguez *et al* 2024 *Quantum Sci. Technol.* **9** 035044

Quantum Science and Technology

PAPER

Hacking coherent-one-way quantum key distribution with present-day technology

Javier Rey-Domínguez[1,2,*], Álvaro Navarrete[2,3,4], Peter van Loock[5] and Marcos Curty[2,3,4]

[1] School of Electronic and Electrical Engineering, Pollard Institute, University of Leeds, Leeds LS2 9JT, United Kingdom
[2] Escuela de Ingeniería de Telecomunicación, Department of Signal Theory and Communications, University of Vigo, Vigo E-36310, Spain
[3] Vigo Quantum Communication Center, University of Vigo, Vigo E-36310, Spain
[4] AtlanTTic Research Center, University of Vigo, Vigo E-36310, Spain
[5] Johannes-Gutenberg University of Mainz, Institute of Physics, Staudingerweg 7, 55128 Mainz, Germany
[*] Author to whom any correspondence should be addressed.

E-mail: j.reydominguez@leeds.ac.uk

Keywords: quantum key distribution, quantum hacking, coherent-one-way

OPEN ACCESS

RECEIVED
9 February 2024

REVISED
3 May 2024

ACCEPTED FOR PUBLICATION
22 May 2024

PUBLISHED
6 June 2024

Abstract

Recent results have shown that the secret-key rate of coherent-one-way (COW) quantum key distribution (QKD) scales quadratically with the system's transmittance, thus rendering this protocol unsuitable for long-distance transmission. This was proven by using a so-called zero-error attack, which relies on an unambiguous state discrimination (USD) measurement. This type of attack allows the eavesdropper to learn the whole secret key without introducing any error. Here, we investigate the feasibility and effectiveness of zero-error attacks against COW QKD with present-day technology. For this, we introduce two practical USD receivers that can be realized with linear passive optical elements, phase-space displacement operations and threshold single-photon detectors. The first receiver is optimal with respect to its success probability, while the second one can impose stronger restrictions on the protocol's performance with faulty eavesdropping equipment. Our findings suggest that zero-error attacks could break the security of COW QKD even assuming realistic experimental conditions.

1. Introduction

Quantum key distribution (QKD) [1–3] has emerged as a cornerstone of quantum cryptography, enabling two remote parties, commonly referred to as Alice and Bob, to share an information-theoretically secure cryptographic key. While QKD networks are currently being deployed worldwide [4–7], QKD still faces certain inherent limitations such as channel loss, which fundamentally restricts the secret-key rate in point-to-point configurations [8, 9], as well as device imperfections that jeopardize the security of practical implementations [2, 10, 11].

Various strategies have been proposed to mitigate these limitations and improve the security, practicality, and performance of QKD systems, including e.g. decoy-state QKD [12–14], measurement-device-independent QKD [15–20], twin-field QKD [21–25], and distributed-phase-reference (DPR) QKD. Among the latter protocols, coherent-one-way (COW) QKD [26–29] has attracted great attention in recent years for its simplicity and its promise to overcome the photon-number-splitting (PNS) attack [30, 31], thus achieving long transmission distances. Indeed, commercial systems implementing the COW protocol have been developed [32] and experimental demonstrations have achieved distances of over 300 km [29]. However, it is important to note that security analyses of COW-QKD that allow for long-distance communications —i.e. that provide lower bounds on the secret-key rate that scale linearly with the channel transmittance η— have been established solely against a restricted class of attacks termed collective attacks [29]. This contrasts with known lower bounds of the order of $O(\eta^2)$ against general attacks [33], a key-rate scaling that has not been improved in recent variants of COW-QKD that disregard its characteristic inter-round interference [34–36].

Crucially, González-Payo *et al* [37] showed very recently that indeed the key rate of COW-QKD scales at most quadratically with the system's transmittance, rendering all long-distance demonstrations of this scheme performed so far insecure against general attacks. This was achieved using a class of intercept-and-resend attacks known as sequential attacks [38–41]. Intercept-and-resend attacks effectively transform the quantum channel into an entanglement-breaking channel, thus preventing the possibility of secret-key generation [42]. Trényi and Curty [43] further refined this strategy by introducing a sequential attack that does not introduce errors in the system. Notably, this latter so-called zero-error attack—which is based on the use of unambiguous state discrimination (USD) measurements [44, 45]— is essentially optimal, in the sense that no other zero-error attack [46] can further limit the maximum achievable distance of COW-QKD.

Importantly, the works in [37, 43] assume an idealized eavesdropper (Eve) with technological capabilities only limited by quantum mechanics. Indeed, this is the standard scenario considered when proving the security of QKD. However, this could be overconservative in certain cases, as the technology required by Eve to implement her attack might not be available in the mid-term future. For instance, the noisy-storage model [47–50] considers the physical assumption that Eve does not have a large reliable quantum memory.

Similarly, in this work we assume that Eve is restricted to use present-day technology and cannot perform perfect quantum operations but employs faulty devices. In this framework, we investigate the practical feasibility of the optimal zero-error attack against COW-QKD proposed in [43]. For this, we introduce two USD receivers that only require off-the-shelf linear passive optical elements, phase-space displacement operations and threshold single-photon detectors (SPDs). Remarkably, the first receiver corresponds to an optimal USD measurement, in the sense that it maximizes the probability of obtaining a conclusive measurement result when distinguishing Alice's signals, but it can only discriminate data signals. The second USD receiver has a lower success probability but can discriminate both data and decoy signals. This latter condition translates into stringent restrictions on the performance of COW QKD with flawed eavesdropping equipment. For both receivers, we derive analytical expressions for the expected values of the key metrics that characterize the COW protocol as a function of the parameters that describe the noise and inefficiencies of Eve's apparatuses. In doing so, we provide a comprehensive framework for evaluating the security of COW-QKD in realistic scenarios. We find that the most critical experimental parameter for the success of Eve's attack seems to be the quality of interference between Alice's weak coherent pulses and her strong light during the displacement operation. Importantly, our results suggest that zero-error attacks are not only a great threat against COW-QKD, but they could break its security with present-day technology.

The paper is structured as follows. In section 2 we introduce the COW-QKD protocol. Next, in section 3, we present the zero-error attack studied in [43]. Then, in section 4 we introduce the first USD receiver, which is able to implement the optimal USD measurement considered in [43]. Besides, in this section we provide a model to incorporate its most relevant imperfections in a practical setting. Next, in section 5 we derive analytical expressions for the expected values of the relevant metrics required to evaluate the security and performance of COW QKD as a function of Eve's faulty equipment, and we investigate the feasibility of the zero-error attack in [43] with present-day technology in section 6. Finally, in section 7 we present our conclusions. The paper also includes several Appendices with additional calculations, which includes the analysis associated to a second USD receiver.

2. COW QKD

The setup for the original COW system [26, 27] is shown in figure 1. In each round, Alice transmits a signal $|\varphi_i\rangle$ to Bob with probability $p_{\mathbf{A}_i}$, where $i \in \{0,1,2\}$. These signals are composed of two optical pulses that could either be in a vacuum state $|0\rangle$ or in a coherent state $|\alpha\rangle$, where $\alpha > 0$. Specifically, the data signals $|\varphi_0\rangle = |0\rangle|\alpha\rangle$ and $|\varphi_1\rangle = |\alpha\rangle|0\rangle$ correspond to the bit values 0 and 1, respectively, and are generated with an equal *a priori* probability $p_{\mathbf{A}_0} = p_{\mathbf{A}_1} = (1-f)/2$, whereas the decoy signal $|\varphi_2\rangle = |\alpha\rangle|\alpha\rangle$ is prepared with probability $p_{\mathbf{A}_2} = f$. Here, temporal sequences of states or signals are represented from right (earlier time) to left (later time).

At Bob's side, an asymmetric beamsplitter with transmittance t_B distributes the incoming signals between the data line and the monitoring line. The former consists of an SPD D_{D} and is used for raw key generation. Specifically, Bob assigns the bit value 0 (1) to a round if D_{D} clicks in the first (second) time slot of that round, and a random bit is assigned in the case of a double click. On the other hand, the monitoring line consists of a Mach-Zehnder interferometer followed by two SPDs D_{M1} and D_{M2}, for constructive and destructive interference, respectively. This line monitors the coherence between adjacent pulses. In particular, the interference is such that two consecutive coherent states $|\alpha\rangle$ cannot trigger D_{M2}, but only D_{M1} (see figure 1).

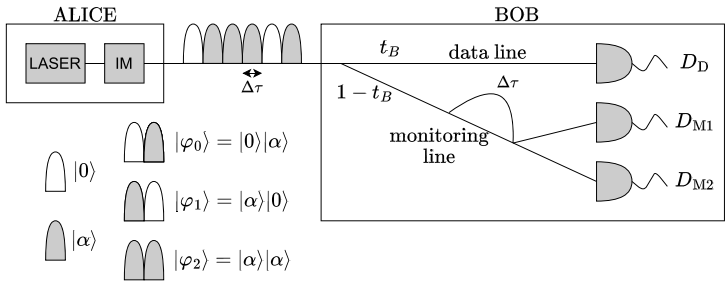

Figure 1. Schematic illustration of the original COW system [26, 27]. Alice sends Bob a random sequence of signals $|\varphi_i\rangle$, with $i \in \{0,1,2\}$. In the figure, we represent coherent (vacuum) states with a grey (white) pulse. Temporal sequences of states or signals are represented from right (earlier time) to left (latter time). Bob uses a beamsplitter with transmittance t_B to distribute the received signals between the data line and the monitoring line. The former is used for key generation, while the latter implements a Mach–Zehnder (MZ) interferometer that measures the coherence between adjacent coherent pulses. This is used to detect eavesdropping. IM refers to the intensity modulator; $\Delta\tau$ is the time delay between consecutive pulses; D_{D}, D_{M1}, and D_{M2} are Bob's SPDs.

Once the quantum communication phase of the protocol ends, Bob publicly announces in which rounds he observed at least one detection click at D_{D}. Then, Alice announces in which of these rounds she prepared a data signal. The bits assigned to this set constitute the sifted key.

Three parameters are specially relevant in COW-QKD. The gain, G, which is the probability that a signal sent by Alice produces at least one detection click in Bob's data line; the quantum bit error rate (QBER) of Alice and Bob's sifted keys; and the visibilities V_s, which quantify the coherence between adjacent coherent pulses, and are computed from the click probabilities in the monitoring line. Specifically, these visibilities are defined as

$$V_s := \frac{p_{\mathrm{M1}|s} - p_{\mathrm{M2}|s}}{p_{\mathrm{M1}|s} + p_{\mathrm{M2}|s}}, \tag{1}$$

where $s \in \{2, 01, 02, 21, 22\}$ represents a sequence of COW signals that contains two adjacent coherent states $|\alpha\rangle$, and $p_{\mathrm{M}i|s}$ is the conditional probability that detector $D_{\mathrm{M}i}$ clicks when such two coherent pulses interfere, given that Alice prepared the sequence s. For example, V_2 characterizes the visibility between the two coherent pulses contained in a decoy signal $|\varphi_2\rangle$, whereas V_{02} characterizes the visibility between the second optical pulse of $|\varphi_2\rangle$ and the first one of $|\varphi_0\rangle$, both of them in the state $|\alpha\rangle$. The other visibilities are interpreted in a similar way. Finally, it is convenient to consider the average visibility, which is given by [29, 37]

$$V_{\mathrm{ave}} := \frac{p_{\mathrm{M1}} - p_{\mathrm{M2}}}{p_{\mathrm{M1}} + p_{\mathrm{M2}}}, \tag{2}$$

where $p_{\mathrm{M}i} = \sum_s p_s p_{\mathrm{M}i|s}$, being p_s the probability that Alice prepares the sequence s.

3. Zero-error attacks against COW-QKD

In a zero-error attack Eve intercepts all of Alice's signals and performs a USD measurement on each of them [37, 43]. We denote by $p_{\mathbf{E}_j|\mathbf{A}_i}$ the probability that Eve obtains the result $\mathbf{E}_j$, given that Alice emits the signal $|\varphi_i\rangle$. Here, $\mathbf{E}_0$, $\mathbf{E}_1$ and $\mathbf{E}_2$ identify the signals $|\varphi_0\rangle$, $|\varphi_1\rangle$ and $|\varphi_2\rangle$, respectively, and $\mathbf{E}_3$ represents an inconclusive outcome. Obviously, in the ideal scenario in which the USD measurement is implemented perfectly, we have by definition that $p_{\mathbf{E}_j|\mathbf{A}_i} = 0\ \forall i \neq j$, with $i,j < 3$.

Next, Eve groups the measured outcomes into blocks for processing them before she sends Bob a regenerated sequence of signals. Precisely, a block of $(k+1)$ signals corresponds to $k \in \{0,1,\ldots,M_{\max}\}$ consecutive conclusive measurement outcomes, followed by an inconclusive measurement result. For example, when $k=0$ the block corresponds to one inconclusive measurement outcome, and Eve sends Bob two vacuum pulses, $|\varphi_{\mathrm{vac}}\rangle = |0\rangle|0\rangle$. If $k=1$, the block has one conclusive measurement outcome, say $|\varphi_i\rangle$, followed by an inconclusive one for which Eve sends Bob the vacuum signal $|\varphi_{\mathrm{vac}}\rangle$. For all cases where $1 < k < M_{\max}$, the interpretation is similar. Finally, if Eve obtains $M_{\max}$ consecutive conclusive measurement outcomes, she ignores the next signal from Alice, and simply treats it as an inconclusive result. The parameter $M_{\max}$ allows Eve to cap the block length, thereby limiting the maximum delay she introduces in the channel. Throughout this paper, we shall use the term *block* to denote the k conclusively measured signals together with the inconclusive measurement outcome, and the term *conclusive-block* when ignoring the latter.

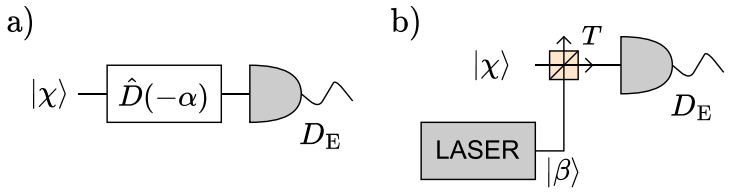

Figure 2. Scheme to implement Eve's optimal USD measurement. (a) Ideal theoretical setup: Eve applies a displacement operator $\hat{D}(-\alpha)$ on each pulse $|\chi\rangle$ of the signal, and subsequently measures the outcome pulses with a threshold detector D_{E}. (b) Experimental implementation of the ideal setup: The displacement operation can be approximated in practice with an asymmetric beamsplitter of transmittance $T \approx 1$, which has at one of its input ports a laser source that emits coherent pulses with amplitude β.

For each conclusive-block, Eve searches for the first and last instances of a vacuum pulse within the block. She then resends to Bob all the optical pulses situated between these two, exactly as she identified them —i.e. if the measurement result with respect to a particular signal is $\mathbf{E}_j$, she resends $|\varphi_j\rangle$— but substitutes the coherent pulses $|\alpha\rangle$ by $|\gamma\rangle$, with $|\gamma|^2 \gg |\alpha|^2$, to increase the detection probability at Bob's side. The remaining pulses within the block that are outside this interval are resent to Bob as vacuum pulses. As already mentioned, the last signal of a block, which corresponds to an inconclusive measurement outcome, is resent as $|\varphi_{\mathrm{vac}}\rangle$. For a more detailed description of the attack, we refer the reader to [43].

Notably, if Eve's equipment is flawless—as considered in [43]— the attack described above introduces no errors on Bob's side, while it reveals Eve full information about the key. This is because Eve's USD measurement ensures that she never misidentifies Alice's signals, while the block processing strategy takes advantage of the fact that vacuum pulses do not introduce errors in the monitoring line.

In particular, [43] showed that, whenever $f/(1-f) \leq 2e^{-|\alpha|^2}$, a regime typically satisfied by practical implementations of COW-QKD, Eve's optimal USD measurement (i.e. the one that maximizes her probability of a conclusive result) satisfies $p_{\mathbf{E}_0|\mathbf{A}_0} = p_{\mathbf{E}_1|\mathbf{A}_1} = 1 - e^{-|\alpha|^2}$ and $p_{\mathbf{E}_0|\mathbf{A}_1} = p_{\mathbf{E}_1|\mathbf{A}_0} = p_{\mathbf{E}_2|\mathbf{A}_i} = 0 \; \forall i \in \{0,1,2\}$.

4. Implementation of zero-error attacks

In this section, we now introduce a linear optics circuit to implement Eve's optimal USD measurement for the zero-error attack described above. It is illustrated in figure 2(a). The input state $|\chi\rangle$, with $\chi \in \{0,\alpha\}$, corresponds to each of the two optical pulses sent by Alice within a signal. That is, for each signal, Eve uses the same scheme in figure 2(a) twice, once per pulse. Precisely, each pulse is displaced according to the transformation $|\chi\rangle \to \hat{D}(-\alpha)|\chi\rangle = |\chi - \alpha\rangle$, where $\hat{D}(x) := e^{x\hat{a}^\dagger - x^*\hat{a}}$ is the displacement operator, and $\hat{a}^\dagger$ and $\hat{a}$ are the creation and annihilation operators, respectively. Finally, the resulting signal is measured with a SPD.

If D_{E} clicks only in the first (second) time slot, the signal is identified as $|\varphi_1\rangle$ ($|\varphi_0\rangle$); in all other cases, the result is inconclusive. This is so because after the optical displacement, the data signals $|\varphi_0\rangle$ and $|\varphi_1\rangle$ are transformed into $|-\alpha\rangle|0\rangle$ and $|0\rangle|-\alpha\rangle$, respectively, while the decoy signal is turned into a vacuum signal $|\varphi_{\mathrm{vac}}\rangle$. Consequently, a single click uniquely identifies a data signal, which occurs with probability $1 - e^{-|\alpha|^2}$, thus matching the optimal probability obtained in [43]. Indeed, this measurement is unable to identify decoy signals, effectively removing $\mathbf{E}_2$ from the POVM set and making $p_{\mathbf{E}_2|\mathbf{A}_i} = 0 \; \forall i$.

In practice, it is well-known that a displacement operation can be approximated with a highly asymmetric beamsplitter of transmittance $T \approx 1$ [51], which has at one of its input ports a coherent state $|\beta\rangle$, as shown in figure 2(b). This scheme transforms $|\chi\rangle$ into $|\sqrt{T}\chi + \sqrt{1-T}\beta\rangle$, so the displacement $\hat{D}(-\alpha)$ can be approximated by setting

$$\beta = -\sqrt{\frac{T}{1-T}}\alpha. \tag{3}$$

Indeed, with this choice, the states $|0\rangle$ and $|\alpha\rangle$ are transformed, respectively, into $|-\sqrt{T}\alpha\rangle$ and $|0\rangle$. This means that the detection probability at D_{E} decreases slightly when Alice transmits $|0\rangle$ (when compared to the case where Eve can use an ideal displacement). However, it remains zero when Alice sends $|\alpha\rangle$. That is, the approximated displacement does not introduce errors.

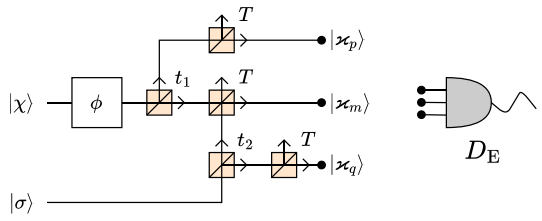

Figure 3. Theoretical model of the setup illustrated in figure 2(b), now incorporating the effect of the main experimental imperfections. $|\chi\rangle$ represents the state of Alice's transmitted pulse, and ϕ represents a phase shift. The beamsplitter transmittances t_1 and t_2 quantify the mode overlap between Alice's pulse and a coherent state $|\sigma\rangle$, whose intensity deviates slightly from that of $|\beta\rangle$, at the beamsplitter with transmittance T (see figure 2(b)). The three outputs of the circuit represent different optical modes in coherent states $|\varkappa_x\rangle$, with $x \in \{m,p,q\}$. The total intensity received in D_{E} is calculated as the sum of the intensities in each of these modes.

4.1. Effect of device imperfections

Now, we investigate the performance of the setup above in a realistic setting, in which we accommodate the most relevant device imperfections. For this, we allow the optical phase of the incoming pulses $|\chi\rangle$ to the beamsplitter in figure 2(b) to be slightly shifted with respect to the laser pulses $|\beta\rangle$. We denote such phase shift by ϕ. Also, we allow for an imperfect mode overlap between the two interfering pulses $|\chi e^{i\phi}\rangle$ and $|\beta\rangle$ at the beamsplitter. To characterize this effect we use the model introduced in [52], which defines two parameters, t_1 and t_2, to quantify, respectively, the fraction of each of Alice's and Eve's input pulses that is properly mode matched at the beamsplitter (see appendix A). In addition, we allow for a non-ideal efficiency η_E and a dark-count probability $p_{\mathrm{d}}^{\mathrm{E}}$ in Eve's SPD D_{E}. Finally, we also account for small intensity fluctuations. For this, we consider a simple model in which the amplitude of $|\beta\rangle$ is slightly deviated from its ideal value (see equation (3)). In particular, we compute its amplitude, which we call now σ to distinguish it from the ideal case, as

$$\sigma = \sqrt{1+\delta}\beta, \tag{4}$$

where the parameter $\delta \in [-1,\infty)$ characterizes the deviation between the ideal and actual intensity.

The schematic of the model that incorporates these imperfections is illustrated in figure 3. First, the incoming pulse $|\chi\rangle$ is phase shifted by ϕ. Then, $|\chi e^{i\phi}\rangle$ and $|\beta\rangle$ are split into two modes each, according to the quantities t_1 and t_2. Two of these optical modes contain the fraction of the input pulses that properly interfere at the beamsplitter with transmittance T. We use the label 'm' to denote the optical mode associated with the output port of this latter beamsplitter that is connected to Eve's detector. The remaining output modes, namely those labeled 'p' and 'q' in figure 3, go through the beamsplitter without interfering. The total optical intensity at D_{E} given that Alice sent the state $|\chi\rangle$, which we shall denote as $\mu_{\mathrm{E}|\chi}$, is thus the sum of the intensities from the three output modes. That is, $\mu_{\mathrm{E}|\chi} = |\varkappa_m|^2 + |\varkappa_p|^2 + |\varkappa_q|^2$, where $|\varkappa_x\rangle$ denotes a coherent state in the output mode x, with $x \in \{m,p,q\}$. We find, therefore, that

$$\mu_{\mathrm{E}|\chi} = T\left[|\chi|^2 + |\alpha|^2(1+\delta) - 2\sqrt{t_1 t_2(1+\delta)}\mathrm{Re}\left\{\chi\alpha^* e^{i\phi}\right\}\right]. \tag{5}$$

Let $p_{\overline{\mathrm{E}}|\chi=0}$ ($p_{\overline{\mathrm{E}}|\chi=\alpha}$) denote the probability of a no-click event in D_{E}, given that $\chi=0$ ($\chi=\alpha$) in that time slot. These probabilities can be computed as $p_{\overline{\mathrm{E}}|\chi} = (1-p_{\mathrm{d}}^{\mathrm{E}})\exp\{-\eta_E\mu_{\mathrm{E}|\chi}\}$. That is,

$$\begin{aligned}
\mu_{\mathrm{E}|\chi=0} &= T|\alpha|^2(1+\delta), \\
\mu_{\mathrm{E}|\chi=\alpha} &= T|\alpha|^2\left[2+\delta-2\sqrt{t_1 t_2(1+\delta)}\cos\phi\right].
\end{aligned} \tag{6}$$

This means, in particular, that the probabilities $p_{\mathbf{E}_j|\mathbf{A}_i}$ can be expressed as follows:

$$\begin{aligned}
p_{\mathbf{E}_0|\mathbf{A}_0} &= p_{\mathbf{E}_1|\mathbf{A}_1} = p_{\overline{\mathrm{E}}|\chi=\alpha}\left(1-p_{\overline{\mathrm{E}}|\chi=0}\right), \\
p_{\mathbf{E}_0|\mathbf{A}_1} &= p_{\mathbf{E}_1|\mathbf{A}_0} = p_{\overline{\mathrm{E}}|\chi=0}\left(1-p_{\overline{\mathrm{E}}|\chi=\alpha}\right), \\
p_{\mathbf{E}_0|\mathbf{A}_2} &= p_{\mathbf{E}_1|\mathbf{A}_2} = p_{\overline{\mathrm{E}}|\chi=\alpha}\left(1-p_{\overline{\mathrm{E}}|\chi=\alpha}\right), \\
p_{\mathbf{E}_2|\mathbf{A}_0} &= p_{\mathbf{E}_2|\mathbf{A}_1} = p_{\mathbf{E}_2|\mathbf{A}_2} = 0, \\
p_{\mathbf{E}_3|\mathbf{A}_i} &= 1-\left(p_{\mathbf{E}_0|\mathbf{A}_i} + p_{\mathbf{E}_1|\mathbf{A}_i}\right) \quad \text{for } i \in \{0,1,2\}.
\end{aligned} \tag{7}$$

5. Performance evaluation

To evaluate the performance of Eve's zero-error attack with current technology, here we derive analytical expressions for the expected values of the three metrics defined in section 2 —namely the gain, the QBER, and the different visibilities V_s— as a function of the parameters that characterize Eve's imperfect operation, as well as the parameters of the protocol.

In the calculations below, we shall consider that γ is sufficiently large to ensure that the signals Eve sends Bob always trigger his detectors unless she sends him a vacuum state, in which case Bob only records a click if a dark count occurs. The dark-count probabilities at Bob's detectors D_{D}, D_{M1}, and D_{M2} are denoted as $p_{\mathrm{d}}^{\mathrm{D}}$, $p_{\mathrm{d}}^{\mathrm{M1}}$, and $p_{\mathrm{d}}^{\mathrm{M2}}$, respectively.

First, we derive the expected gain G. Then, we present the expressions to compute the QBER and the visibilities V_s. The full derivation of these latter parameters can be found in appendix B.

5.1. Gain

The gain is defined as the probability that Bob observes at least one click in D_{D} in a round. This quantity can be written as $G = N_{\mathrm{clk}}/N_{\mathrm{sig}}$, where N_{clk} is the average number of signals within a block that produce a click—single or double—at Bob's side, and N_{sig} is the average number of signals within a block.

Let $p_{\mathbf{E}_c}$ be the probability that Eve's USD measurement is conclusive, computed as

$$p_{\mathbf{E}_c} = \sum_{i,j=0}^{2} p_{\mathbf{A}_j} p_{\mathbf{E}_i|\mathbf{A}_j}. \tag{8}$$

Also, let $p_{\mathrm{cb}}(k)$ be the probability that Eve processes a conclusive-block of length k, which is given by [43]

$$p_{\mathrm{cb}}(k) = \begin{cases} p_{\mathbf{E}_c}^{k}\left(1 - p_{\mathbf{E}_c}\right) & \text{when } 0 \leq k < M_{\max}, \\ p_{\mathbf{E}_c}^{M_{\max}} & \text{when } k = M_{\max}, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Then, we have that $N_{\mathrm{sig}} = \sum_{k=0}^{M_{\max}} p_{\mathrm{cb}}(k)(k+1)$, and therefore [43]

$$N_{\mathrm{sig}} = \frac{1 - p_{\mathbf{E}_c}^{M_{\max}+1}}{1 - p_{\mathbf{E}_c}}. \tag{10}$$

The average N_{clk} admits a similar decomposition in terms of the length of a block. In particular,

$$N_{\mathrm{clk}} = \sum_{k=0}^{M_{\max}} p_{\mathrm{cb}}(k)\left[n_{\mathrm{clk}}(k) + p_{\mathrm{clk}}^{\mathrm{last}}(k)\right], \tag{11}$$

where $n_{\mathrm{clk}}(k)$ is the average number of signals within a conclusive-block of length k in which Bob observes at least one click in D_{D}, and $p_{\mathrm{clk}}^{\mathrm{last}}(k)$ is the probability that a click occurs in the last vacuum signal $|\varphi_{\mathrm{vac}}\rangle$ of the block, corresponding to the inconclusive measurement outcome that happened after a conclusive-block of length k. If we define $p_{\mathrm{clk}|\mathbf{B}_v}$ as the probability that Bob observes at least one click in a round where he receives a vacuum signal from Eve, then it is clear that

$$p_{\mathrm{clk}}^{\mathrm{last}}(k) = p_{\mathrm{clk}|\mathbf{B}_v} = 1 - \left(1 - p_{\mathrm{d}}^{\mathrm{D}}\right)^2, \tag{12}$$

for $k \in [0, M_{\max}]$.

Now, let $p_{\mathbf{E}_j|\mathbf{E}_c}$ be the probability that Eve obtains the outcome $\mathbf{E}_j$, given that her measurement was conclusive. Its value is given by

$$p_{\mathbf{E}_j|\mathbf{E}_c} = \frac{1}{p_{\mathbf{E}_c}} \sum_{i=0}^{2} p_{\mathbf{A}_i} p_{\mathbf{E}_j|\mathbf{A}_i}. \tag{13}$$

Then, one can express

$$n_{\mathrm{clk}}(k) = \sum_{i=0}^{2} p_{\mathbf{E}_i|\mathbf{E}_c} n_{\mathrm{clk}}(k|i), \tag{14}$$

where $n_{\text{clk}}(k|i)$ is defined as the average number of signals where a click occurs within a conclusive-block of length k, given that the first signal of the block was identified by Eve as $|\varphi_i\rangle$. Importantly, we note that the quantities $n_{\text{clk}}(k|i)$ admit recursive formulations [43], which one can solve to compute $n_{\text{clk}}(k)$.

To illustrate this, let us focus on $n_{\text{clk}}(k|0)$, i.e. the case where the first signal of the block is identified by Eve as $|\varphi_0\rangle$. According to her block-processing strategy, Eve translates this first signal into vacuum, since the first optical pulse of this signal is a coherent state $|\alpha\rangle$. Consequently, we can disregard it and consider the reduced conclusive-block of length $(k-1)$. Importantly, according to Eve's processing strategy, this also implies that the first signal in the resulting truncated conclusive-block will be resent exactly as identified by her. This is so because it is preceded by a vacuum pulse, given by the second optical pulse of $|\varphi_0\rangle$. Then, it only remains to find the last signal of the conclusive-block that is not resent as vacuum.

There are three possibilities, which depend on the last signal identified by Eve. If this signal is $|\varphi_0\rangle$, then it directly becomes the last non-vacuum signal of the block, and thus $(k-1)$ non-vacuum signals will arrive at D_{D}. If the last signal is $|\varphi_1\rangle$, it is translated into vacuum, and the preceding $(k-2)$ non-vacuum signals will arrive at D_{D}. Finally, if the last signal is $|\varphi_2\rangle$, this signal is also translated into vacuum, and, by definition, Bob will observe, on average, $n_{\text{clk}}(k-1|0)$ clicks in D_{D}. Putting all this together, one can write $n_{\text{clk}}(k-1|i)$ as

$$\begin{aligned} n_{\text{clk}}(k|0) &= p_{\mathbf{E}_0|\mathbf{E}_c}\left(k-1+p_{\text{clk}|\mathbf{B}_v}\right)+p_{\mathbf{E}_1|\mathbf{E}_c}\left(k-2+2p_{\text{clk}|\mathbf{B}_v}\right)+p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\text{clk}}(k-1|0)+p_{\text{clk}|\mathbf{B}_v}\right], \\ n_{\text{clk}}(k|1) &= p_{\mathbf{E}_0|\mathbf{E}_c}k+p_{\mathbf{E}_1|\mathbf{E}_c}\left(k-1+p_{\text{clk}|\mathbf{B}_v}\right)+p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\text{clk}}(k-1|1)+p_{\text{clk}|\mathbf{B}_v}\right], \\ n_{\text{clk}}(k|2) &= n_{\text{clk}}(k-1)+p_{\text{clk}|\mathbf{B}_v}. \end{aligned} \tag{15}$$

Moreover, the starting points for the previous recursions are $n_{\text{clk}}(1|0)=n_{\text{clk}}(1|1)=p_{\text{clk}|\mathbf{B}_v}$ and $n_{\text{clk}}(0)=0$. With this, one can solve the recursion and obtain

$$n_{\text{clk}}(k)=k+kR^k\left(1-p_{\text{d}}^{\text{D}}\right)^2-\frac{1+R}{1-R}\left(1-R^k\right)\left(1-p_{\text{d}}^{\text{D}}\right)^2, \tag{16}$$

where we have used R as a shorthand for the recursion factor,

$$R\equiv p_{\mathbf{E}_2|\mathbf{E}_c}. \tag{17}$$

5.2. QBER

Next, we analyze the QBER. This quantity can be written as $\text{QBER}=p_{\text{err}}/p_{\text{key}}$, where p_{key} is the probability that Bob distills a key bit in a given round, and p_{err} the probability that he distills an erroneous key bit. We have that, asymptotically, $p_{\text{key}}=N_{\text{key}}/N_{\text{sig}}$ and $p_{\text{err}}=N_{\text{err}}/N_{\text{sig}}$, where N_{key} (N_{err}) is the average number of key bits (erroneous key bits) distilled by Bob from a block sent by Eve. Therefore, one can rewrite the error rate as [43]

$$\text{QBER}=\frac{N_{\text{err}}}{N_{\text{key}}}. \tag{18}$$

To determine N_{key} and N_{err} we decompose them according to the length of the block processed by Eve. For this, we define $n_{\text{key}}(k)$ ($n_{\text{err}}(\text{k})$) as the average number of key bits (erroneous key bits) distilled from a conclusive-block of length k, and $p_{\text{key}}^{\text{last}}(k)$ ($p_{\text{err}}^{\text{last}}(k)$) as the probability that Bob distills a key bit (erroneous key bit) from the vacuum signal $|\varphi_{\text{vac}}\rangle$ that is sent after a conclusive-block of length k, due to an inconclusive result. Then, we have that

$$\begin{aligned} N_{\text{key}} &= \sum_{k=0}^{M_{\max}} p_{\text{cb}}(k)\left[n_{\text{key}}(k)+p_{\text{key}}^{\text{last}}(k)\right], \\ N_{\text{err}} &= \sum_{k=0}^{M_{\max}} p_{\text{cb}}(k)\left[n_{\text{err}}(k)+p_{\text{err}}^{\text{last}}(k)\right]. \end{aligned} \tag{19}$$

All that remains is to calculate the values of $n_{\text{key}}(k)$, $n_{\text{err}}(\text{k})$, $p_{\text{key}}^{\text{last}}(k)$ and $p_{\text{err}}^{\text{last}}(k)$ that appear in the previous equations. For this, let us define $p_{\text{err}|\mathbf{B}_v}$ as the probability that Bob obtains an incorrect bit from a vacuum signal sent by Eve, given that he distills a bit that round. Its value is given by

$$p_{\text{err}|\mathbf{B}_v}=\frac{p_{\text{d}}^{\text{D}}\left(2-p_{\text{d}}^{\text{D}}\right)}{2}. \tag{20}$$

Then, it can be shown (see appendix B) that $n_{\text{key}}(k)$ has the form

$$n_{\text{key}}(k) = \frac{p_{\mathbf{A}_0}}{p_{\mathbf{E}_c}}\Big\{k\left(2 - p_{\mathbf{E}_3|\mathbf{A}_0} - p_{\mathbf{E}_3|\mathbf{A}_1}\right) + kR^{k-1}\left(1-p_{\text{d}}^{\text{D}}\right)^2\left(p_{\mathbf{E}_2|\mathbf{A}_0} + p_{\mathbf{E}_2|\mathbf{A}_1}\right) - \frac{1-R^k}{1-R}\left(1-p_{\text{d}}^{\text{D}}\right)^2\left(2 - p_{\mathbf{E}_3|\mathbf{A}_0} - p_{\mathbf{E}_3|\mathbf{A}_1} + p_{\mathbf{E}_2|\mathbf{A}_0} + p_{\mathbf{E}_2|\mathbf{A}_1}\right)\Big\}; \tag{21}$$

the parameter $p_{\text{key}}^{\text{last}}$ is given by

$$p_{\text{key}}^{\text{last}}(k) = \begin{cases} \frac{p_{\mathbf{A}_0} p_{\text{clk}|\mathbf{B}_\text{v}}\left(p_{\mathbf{E}_3|\mathbf{A}_0} + p_{\mathbf{E}_3|\mathbf{A}_1}\right)}{1-p_{\mathbf{E}_c}} & \text{if } 0 \leq k < M_{\max}, \\ 2p_{\mathbf{A}_0} p_{\text{clk}|\mathbf{B}_\text{v}} & \text{if } k = M_{\max}, \end{cases} \tag{22}$$

the parameter $n_{\text{err}}(k)$ has the form

$$\begin{aligned} n_{\text{err}}(k) = \frac{p_{\mathbf{A}_0}}{2p_{\mathbf{E}_c}}\Big\{k\Big[\left(p_{\mathbf{E}_0|\mathbf{A}_0} + p_{\mathbf{E}_1|\mathbf{A}_1}\right)p_{\text{d}}^{\text{D}} + p_{\mathbf{E}_2|\mathbf{A}_0} + p_{\mathbf{E}_2|\mathbf{A}_1} + \left(p_{\mathbf{E}_0|\mathbf{A}_1} + p_{\mathbf{E}_1|\mathbf{A}_0}\right)\left(2-p_{\text{d}}^{\text{D}}\right)\Big] \\ + kR^{k-1}\left(p_{\mathbf{E}_2|\mathbf{A}_0} + p_{\mathbf{E}_2|\mathbf{A}_1}\right)\left(1-p_{\text{d}}^{\text{D}}\right)^2 + \frac{1-R^k}{1-R}\Big[\left(p_{\mathbf{E}_0|\mathbf{A}_0} + p_{\mathbf{E}_1|\mathbf{A}_1}\right)p_{\text{d}}^{\text{D}} \\ - \left(p_{\mathbf{E}_0|\mathbf{A}_1} + p_{\mathbf{E}_1|\mathbf{A}_0}\right)\left(2-p_{\text{d}}^{\text{D}}\right) - 2\left(p_{\mathbf{E}_2|\mathbf{A}_0} + p_{\mathbf{E}_2|\mathbf{A}_1}\right)\left(1-p_{\text{d}}^{\text{D}}\right)\Big]\left(1-p_{\text{d}}^{\text{D}}\right)\Big\}, \end{aligned} \tag{23}$$

and the parameter $p_{\text{err}}^{\text{last}}(k)$ can be expressed as

$$p_{\text{err}}^{\text{last}}(k) = \begin{cases} \frac{p_{\mathbf{A}_0} p_{\text{err}|\mathbf{B}_\text{v}}\left(p_{\mathbf{E}_3|\mathbf{A}_0} + p_{\mathbf{E}_3|\mathbf{A}_1}\right)}{1-p_{\mathbf{E}_c}} & \text{if } 0 \leq k < M_{\max}, \\ 2p_{\mathbf{A}_0} p_{\text{err}|\mathbf{B}_\text{v}} & \text{if } k = M_{\max}. \end{cases} \tag{24}$$

5.3. Visibilities

To calculate the resulting visibilities V_s, we start by expressing the probabilities that appear in equation (1) as $p_{MX|s} = p_{MX,s}/p_s$, with $X \in \{1,2\}$. The joint probabilities $p_{MX,s}$ can be written as $p_{MX,s} = N_{MX,s}/N_{\text{sig}}$. Here, $N_{MX,s}$ is the average number of times in which the sequence s appears within a block and a click is registered in D_{MX} during the time slot associated with the interference of the two intermediate coherent pulses of the sequence. This means that V_s can be expressed as

$$V_s = \frac{N_{\text{M1},s} - N_{\text{M2},s}}{N_{\text{M1},s} + N_{\text{M2},s}}. \tag{25}$$

Let us now start with the case $s = 2$, since this is the only one that considers interference between pulses within the same signal. By applying analogous reasoning to that used to derive equations (11) and (19), it can be shown that $N_{MX,2}$ can be expressed as

$$N_{MX,2} = \sum_{k=0}^{M_{\max}} p_{\text{cb}}(k)\left[n_{MX,2}(k) + p_{MX,2}^{\text{last}}(k)\right], \tag{26}$$

where $n_{MX,2}(k)$ represents the average number of signals within a conclusive block of length k in which Alice sends $|\varphi_2\rangle$ and D_{MX} clicks, and $p_{MX,2}^{\text{last}}(k)$ is the probability that this event occurs in the last signal of the full block. Precisely, it is shown in appendix B that these quantities can be written as

$$\begin{aligned} n_{\text{M1},2}(k) &= \frac{p_{\mathbf{A}_2}}{p_{\mathbf{E}_c}}\left\{k\left(1-p_{\mathbf{E}_3|\mathbf{A}_2}\right) + kR^{k-1}p_{\mathbf{E}_2|\mathbf{A}_2}\left(1-p_{\text{d}}^{\text{M1}}\right) - \frac{1-R^k}{1-R}\left(1 + p_{\mathbf{E}_2|\mathbf{A}_2} - p_{\mathbf{E}_3|\mathbf{A}_2}\right)\left(1-p_{\text{d}}^{\text{M1}}\right)\right\}, \\ n_{\text{M2},2}(k) &= \frac{p_{\mathbf{A}_2}}{p_{\mathbf{E}_c}}\Big\{k\left(p_{\mathbf{E}_0|\mathbf{A}_2} + p_{\mathbf{E}_1|\mathbf{A}_2}\right) - kR^{k-1}p_{\mathbf{E}_2|\mathbf{A}_2}p_{\text{d}}^{\text{M2}} - \frac{1-R^k}{1-R}\Big[\left(p_{\mathbf{E}_0|\mathbf{A}_2} + p_{\mathbf{E}_1|\mathbf{A}_2}\right)\left(1-p_{\text{d}}^{\text{M2}}\right) \\ &\quad - 2p_{\mathbf{E}_2|\mathbf{A}_2}p_{\text{d}}^{\text{M2}}\Big]\Big\}, \end{aligned} \tag{27}$$

and

$$p_{MX,2}^{\text{last}}(k) = \begin{cases} \frac{p_{\mathbf{A}_2} p_{\mathbf{E}_3|\mathbf{A}_2} p_{\text{d}}^{MX}}{1-p_{\mathbf{E}_c}} & \text{if } 0 \leq k < M_{\max}, \\ p_{\mathbf{A}_2} p_{\text{d}}^{MX} & \text{if } k = M_{\max}, \end{cases} \tag{28}$$

where R is given by equation (17).

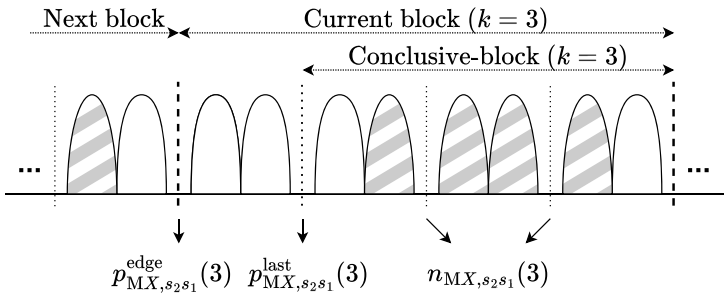

Figure 4. Schematic illustration of the different interference events at Bob's side considered within a block, and the quantities involved in the computation of N_{MX,s_2s_1} (see equation (29)). For simplicity, this figure shows a block with $k=3$. White pulses represent vacuum states, $|0\rangle$, sent to Bob by Eve, whereas the stripped pulses may be either $|0\rangle$ or $|\alpha\rangle$. Being a valid block, it has two vacuum pulses at the conclusive-block's surrounding borders, as well as a vacuum signal $|\varphi_{\text{vac}}\rangle$ at the end of the block.

We now consider the remaining visibilities, where we denote the two-signal interference sequence as $s \equiv s_2 s_1$, meaning Alice first prepares $|\varphi_{s_1}\rangle$ and subsequently prepares $|\varphi_{s_2}\rangle$. Importantly, a subtle nuance must be considered in the definition of the averages N_{MX,s_2s_1}. Since we are dealing with the interference between adjacent pulses in consecutive rounds, these rounds may belong to different blocks. If this happens, we will consider that the observed clicks are attributed to the first block. This is an arbitrary decision, but it does not impact the final result.

Once again, we express the averages under analysis as a decomposition over the block length. For this, we define $n_{MX,s_2s_1}(k)$ as the average number of times within a conclusive-block of length k where Alice sends $s = s_2 s_1$ and a click is registered in D_{MX} in the time slot associated with the interference between the two signals. Those events in which one of the signals of the sequence s does not belong to the considered conclusive-block are not accounted in $n_{MX,s_2s_1}(k)$. Moreover, we denote as $p^{\text{last}}_{MX,s_2s_1}$ the probability that Alice prepares $|\varphi_{s_1}\rangle$ in the last round of a conclusive-block of length k and $|\varphi_{s_2}\rangle$ in the next round, and Bob observes a click in D_{MX} in the time slot associated with the interference of these two signals. Similarly, we define $p^{\text{edge}}_{MX,s_2s_1}$ as the probability that Alice prepares $|\varphi_{s_1}\rangle$ in the last round of the full block of $k+1$ signals, $|\varphi_{s_2}\rangle$ in the first round of the next block, and Bob observed a click in D_{MX} in the time slot associated with the interference of these two signals. With these definitions, which are illustrated in figure 4, we can write N_{MX,s_2s_1} as

$$N_{MX,s_2s_1} = \sum_{k=0}^{M_{\max}} p_{\text{cb}}(k) \left[n_{MX,s_2s_1}(k) + p^{\text{last}}_{MX,s_2s_1}(k) + p^{\text{edge}}_{MX,s_2s_1}(k) \right]. \tag{29}$$

Finally, let $p_{\mathbf{A}_i|\mathbf{E}_c}$ ($p_{\mathbf{A}_i|\mathbf{E}_3}$) be the probability that Alice sent the signal $|\varphi_i\rangle$, given that Eve's measurement was conclusive (inconclusive). These values can be written as

$$\begin{aligned} p_{\mathbf{A}_i|\mathbf{E}_c} &= \frac{p_{\mathbf{A}_i}\left(1 - p_{\mathbf{E}_3|\mathbf{A}_i}\right)}{p_{\mathbf{E}_c}}, \\ p_{\mathbf{A}_i|\mathbf{E}_3} &= \frac{p_{\mathbf{A}_i} p_{\mathbf{E}_3|\mathbf{A}_i}}{1 - p_{\mathbf{E}_c}}, \end{aligned} \tag{30}$$

for $i \in \{0,1,2\}$. Then, appendix B shows that

$$\begin{aligned} n_{\text{M1},s_2s_1}(k) = \frac{p_{\mathbf{A}_{s_1}} p_{\mathbf{A}_{s_2}}}{p^2_{\mathbf{E}_c}} \Big\{ & (k-1)\Big[\left(1 - p_{\mathbf{E}_3|\mathbf{A}_{s_1}}\right)\left(1 - p_{\mathbf{E}_3|\mathbf{A}_{s_2}}\right) - p_{\mathbf{E}_0|\mathbf{A}_{s_1}} p_{\mathbf{E}_1|\mathbf{A}_{s_2}} \left(1 - p^{\text{M1}}_{\text{d}}\right) \Big] \\ & + (k-1) R^{k-2} \left(1 - p^{\text{M1}}_{\text{d}}\right) p_{\mathbf{E}_2|\mathbf{A}_{s_1}} p_{\mathbf{E}_2|\mathbf{A}_{s_2}} - \frac{1 - R^{k-1}}{1 - R} \left(1 - p^{\text{M1}}_{\text{d}}\right) \Big[\left(1 - p_{\mathbf{E}_3|\mathbf{A}_{s_1}}\right) p_{\mathbf{E}_2|\mathbf{A}_{s_2}} \\ & + p_{\mathbf{E}_2|\mathbf{A}_{s_1}} \left(1 - p_{\mathbf{E}_3|\mathbf{A}_{s_2}}\right) \Big] \Big\}, \end{aligned} \tag{31}$$

and

$$n_{\mathrm{M2},s_2s_1}(k)=\frac{p_{\mathbf{A}_{s_1}}p_{\mathbf{A}_{s_2}}}{p_{\mathbf{E}_c}^2}\Big\{(k-1)\Big[p_{\mathbf{E}_0|\mathbf{A}_{s_1}}p_{\mathbf{E}_0|\mathbf{A}_{s_2}}+p_{\mathbf{E}_0|\mathbf{A}_{s_1}}p_{\mathbf{E}_2|\mathbf{A}_{s_2}}+p_{\mathbf{E}_1|\mathbf{A}_{s_1}}p_{\mathbf{E}_1|\mathbf{A}_{s_2}}+p_{\mathbf{E}_2|\mathbf{A}_{s_1}}p_{\mathbf{E}_1|\mathbf{A}_{s_2}}$$
$$+p_{\mathrm{d}}^{\mathrm{M2}}\Big(p_{\mathbf{E}_0|\mathbf{A}_{s_1}}p_{\mathbf{E}_1|\mathbf{A}_{s_2}}+p_{\mathbf{E}_1|\mathbf{A}_{s_1}}p_{\mathbf{E}_0|\mathbf{A}_{s_2}}+p_{\mathbf{E}_1|\mathbf{A}_{s_1}}p_{\mathbf{E}_2|\mathbf{A}_{s_2}}+p_{\mathbf{E}_2|\mathbf{A}_{s_1}}p_{\mathbf{E}_0|\mathbf{A}_{s_2}}+p_{\mathbf{E}_2|\mathbf{A}_{s_1}}p_{\mathbf{E}_2|\mathbf{A}_{s_2}}\Big)\Big] \tag{32}$$
$$-\frac{1-R^{k-1}}{1-R}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\left(p_{\mathbf{E}_0|\mathbf{A}_{s_1}}p_{\mathbf{E}_2|\mathbf{A}_{s_2}}+p_{\mathbf{E}_2|\mathbf{A}_{s_1}}p_{\mathbf{E}_1|\mathbf{A}_{s_2}}\right)\Big\},$$

while the parameters $p^{\mathrm{last}}_{MX,s_2s_1}(k)$ and $p^{\mathrm{edge}}_{MX,s_2s_1}(k)$, for $X\in\{1,2\}$, are given by

$$p^{\mathrm{last}}_{MX,s_2s_1}(k)=\begin{cases}0 & \text{if } k=0,\\ p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_3}p_{\mathrm{d}}^{MX} & \text{if } 0<k<M_{\max},\\ p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}}p_{\mathrm{d}}^{MX} & \text{if } k=M_{\max},\end{cases} \tag{33}$$
$$p^{\mathrm{edge}}_{MX,s_2s_1}(k)=\begin{cases}p_{\mathbf{A}_{s_1}|\mathbf{E}_3}p_{\mathbf{A}_{s_2}}p_{\mathrm{d}}^{MX} & \text{if } 0\leq k<M_{\max},\\ p_{\mathbf{A}_{s_1}}p_{\mathbf{A}_{s_2}}p_{\mathrm{d}}^{MX} & \text{if } k=M_{\max}.\end{cases}$$

Finally, we note that V_{ave}, as defined in equation (2), can be directly obtained from the previous averages $N_{MX,s}$ as follows

$$V_{\mathrm{ave}}=\frac{\sum_s N_{\mathrm{M1},s}-\sum_s N_{\mathrm{M2},s}}{\sum_s N_{\mathrm{M1},s}+\sum_s N_{\mathrm{M2},s}} \tag{34}$$

where the sums run for $s\in\{2,01,02,21,22\}$.

6. Simulation results

In this section we now evaluate the feasibility of the zero-error attack analyzed above in realistic conditions. We shall denote the attack that uses the optimal USD measurement discussed in section 4 as USD1. Here, we shall include as well the results derived in appendix C regarding an alternative, suboptimal USD measurement for Eve, which is able to discriminate not only data signals but also decoy signals, although its success probability is lower than that of USD1. We shall denote this second strategy as USD2. We refer the readers to appendix C for more details. As it is shown below, USD2 can outperform USD1 in the presence of imperfections.

We consider that Eve's attack is successful if she can keep the resulting QBER and visibilities within certain acceptance intervals. For illustration purposes, we shall use $\mathrm{QBER}^{\mathrm{th}}=0.05$ and $V^{\mathrm{th}}=0.95$ as the threshold values defining these acceptance intervals, i.e. the QBER and the visibilities V_s must satisfy $\mathrm{QBER}\leq\mathrm{QBER}^{\mathrm{th}}$ and $V_s\geq V^{\mathrm{th}}$ for the attack to be successful. We remark, however, that these values are just an example chosen for continuity with previous studies [37], which in turn selected them to reflect the metrics attainable by state-of-the-art experiments [28, 29]. In any case, our analysis can be applied to any other threshold values used to calculate the secret-key rate by considering a specific security proof. Importantly, some commercial systems only check the average visibility V_{ave} out of all the visibilities [29, 32]. Below we show that this provides Eve a crucial advantage for her attack.

Figure 5 illustrates the resulting gain G and QBER as a function of an error parameter ε, which is directly related to the quality of the mode overlap at Eve's beamsplitters. Specifically, we set $t_1=t_2=1-\varepsilon$ for USD1, and $t_1=t_2=t_3=t_4=1-\varepsilon$ for USD2 (see appendix C). Here we focus on the mode overlap because our results suggest that this is the main limiting experimental factor in Eve's attack, while its effectiveness varies only slightly when the experimental parameters that model other imperfections are changed over realistic ranges of values (see appendix D for further details). In particular, for the simulations in figure 5, we assign to other imperfections the values given in table 1. For simplicity, we consider the same dark-count probability p_{d} for all detectors, i.e. $p_{\mathrm{d}}^{\mathrm{E}}=p_{\mathrm{d}}^{\mathrm{D}}=p_{\mathrm{d}}^{MX}=p_{\mathrm{d}}$, with $X\in\{1,2\}$. Besides, we set the parameters of the COW protocol, namely α and f, to typical values, also shown in table 1, and we fix $M_{\max}=10$, as we observe essentially no improvement of the metrics when increasing this value, given that the intensity of Alice's pulses is kept within practical margins (further details are discussed in appendix D).

Remarkably, figure 5 shows that the gain G grows with ε. To understand this, we note that, for low values of $|\alpha|^2$ and small imperfections, both USD measurements have a small probability of being conclusive. However, as imperfections escalate, the probability of a conclusive measurement increases slightly due to a slight rise in the click probability of Eve's detectors. As expected, USD1 allows for a higher gain than USD2,

Table 1. Protocol and experimental parameters used in the simulations. We consider the same dark-count probability p_d for all detectors, i.e. $p_d^E = p_d^D = p_d^{MX} = p_d$.

f	$\lvert\alpha\rvert^2$	$M_{\max}$	ϕ	T	δ	η_E	p_d
0.155	0.1	10	1°	0.99	0.05	0.6	10^{-7}

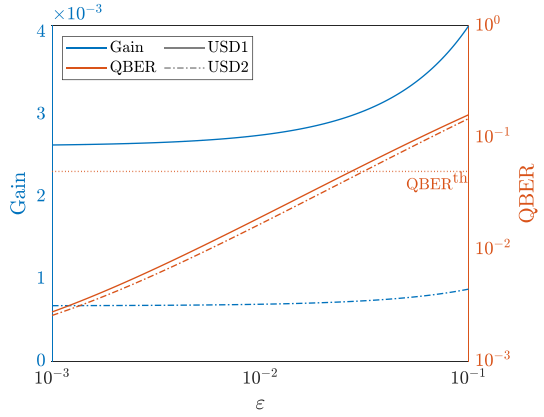

Figure 5. Resulting gain G and QBER for a COW QKD system in the presence of Eve's attack as a function of the error parameter ε. This parameter models Eve's imperfect mode overlap in her measurement implementation. Solid lines correspond to USD1, while dashed lines correspond to USD2. For the simulations we considered the parameters given in table 1.

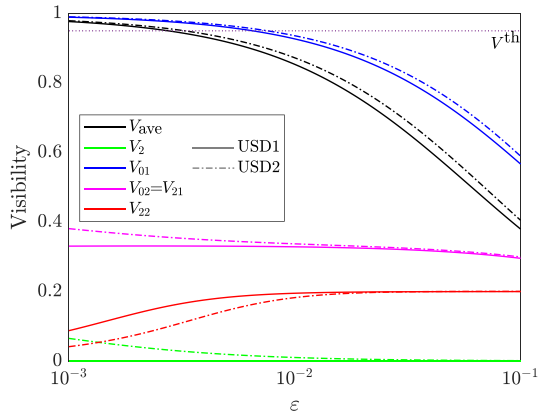

Figure 6. Resulting visibilities for a COW QKD system in the presence of Eve's attack as a function of the error parameter ε. This parameter models Eve's imperfect mode overlap in her measurement implementation. Solid lines correspond to USD1, while dashed lines correspond to USD2. For the simulations we consider the parameters given in table 1.

as the former has been designed to maximize the probability of identifying data signals. As for the QBER, the figure shows that Eve can keep it below the chosen threshold value of 0.05 for reasonably high values of ε.

The visibilities are illustrated in figure 6. We observe that those corresponding to sequences containing a decoy signal are well below the threshold Vth, being this particularly evident for V_2, which is in fact always zero for USD1. This is expected, as USD1 is unable to identify $|\varphi_2\rangle$. Indeed, when receiving $|\varphi_2\rangle$, Eve will mostly resend vacuum signals to Bob. Certainly, due to measurement errors, she could occasionally misidentify Alice's signal and resend Bob $|\varphi_0\rangle$ or $|\varphi_1\rangle$. However, these signals can trigger both detectors in Bob's monitoring line with equal probability, and consequently do not increase the visibility. Needless to say, dark counts cannot increase the expected visibility either, as they occur with equal probability in both detectors. Interestingly, even for USD2, the visibility V_2 is notably low. The reason for this behavior is twofold: first, the probability that Eve identifies $|\varphi_2\rangle$ is relatively low. Second, the probability of obtaining a conclusive measurement outcome is also low for the considered value of $|\alpha|^2$, so most blocks processed by Eve are expected to be short. Therefore, it is likely that those few $|\varphi_2\rangle$ that are correctly identified are located at the edge of a block, where they are consequently erased due to Eve's block-processing strategy.

Similar arguments apply to V_{22}, V_{21}, and V_{02} (which is equal to V_{21} due to the symmetry of the setups). In these cases, however, the visibilities are non-zero even for USD1. This is because their corresponding sequences are occasionally translated by Eve to the sequence '01', which positively contributes to the visibility

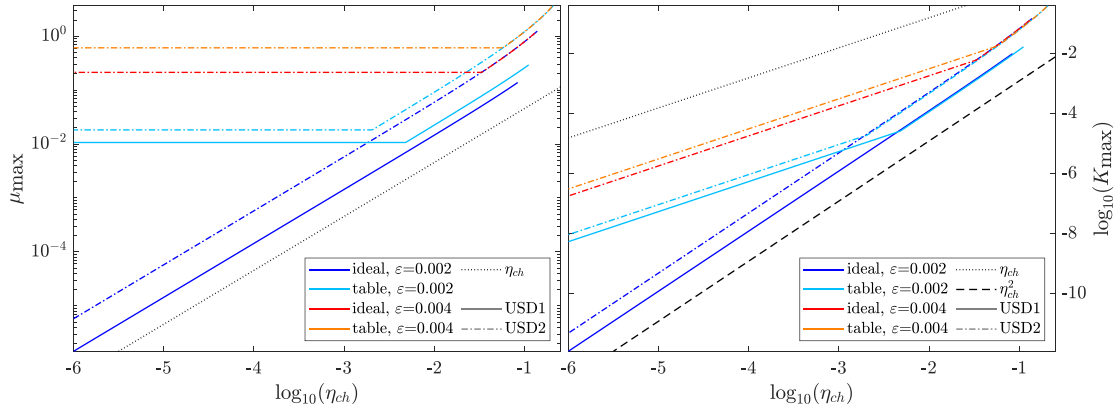

Figure 7. Maximum intensity that Alice can use while still being able to detect Eve's attack, $\mu_{\max}$, and upper bound of the key rate, $K_{\max}$, for different channel transmittances. The lines labeled as *ideal* in the legend were computed by considering no imperfections in Eve's measurement besides the mode mismatch characterized by ε, while those labeled *table* used the parameters in table 1. The dotted (dashed) black line is added as a reference to show linear (quadratic) scaling with $\eta_{\rm ch}$.

(see appendix E for further details). Remarkably, we note that USD2 performs better than USD1 for most visibilities in this scenario. This is because the probability of erroneously identifying $|\varphi_0\rangle$ as $|\varphi_1\rangle$ (or viceversa), given that Eve's measurement is conclusive, is smaller in USD2. For the same reason, V_{01} is also larger for USD2. This is relevant because Alice typically transmits the sequence '01' much more frequently than the other sequences, as f is usually set to a small value to increase the number of data rounds. As a consequence, V_{01}, the highest visibility, is also the largest contributor to the average visibility $V_{\rm ave}$, which thus remains relatively high for reasonably small values of ε.

To compare the performance of Eve's attack with respect to prior studies [37, 43], we compute here a simple upper bound on the secret-key rate. With this in mind, let us conveniently refer to Eve's attack as *undetectable* if the following two conditions are met. Firstly, both the QBER and the average visibility $V_{\rm ave}$ observed for the attacked system must fall within their corresponding acceptance regions, i.e. $\mathrm{QBER} < \mathrm{QBER}^{\rm th}$ and $V_{\rm ave} > V^{\rm th}$. Secondly, the gain G of the attacked system must equal or exceed that expected from a legitimate system, which we will denote as $G^{\bar{\rm a}}$ to indicate that no attack is launched. In particular, in the simulations we consider a typical lossy channel model for which [43]

$$G^{\bar{\rm a}} = 1 - \left[(1-f)\, e^{-\eta_B \eta_{\rm ch} t_B |\alpha|^2} + f e^{-2\eta_B \eta_{\rm ch} t_B |\alpha|^2}\right] \left(1 - p_{\rm d}^{\rm D}\right)^2, \tag{35}$$

where η_B is the efficiency of Bob's detectors, t_B is transmittance of the beamsplitter used by Bob to separate between data and monitoring line (see figure 1), $\eta_{\rm ch} = 10^{-\alpha_{\rm ch} d/10}$ is the channel transmittance, d is the channel distance (in km), and $\alpha_{\rm ch}$ is the attenuation coefficient (in dB/km). Importantly, if for a given $\eta_{\rm ch}$ Eve's attack meets the previous two conditions —i.e. if it is undetectable—it immediately follows that no secret key can be distilled by Alice and Bob based on the observed metrics. Indeed, the secret-key rate K can be simply upper bounded as [37, 43]

$$K < (1-f)\, \eta_{\rm ch} \eta_B \mu_{\max} \equiv K_{\max}, \tag{36}$$

where $\mu_{\max}$ is the maximum value of μ for which Eve cannot perform an undetectable attack. We remark that, in general, $\mu_{\max}$ depends on $\eta_{\rm ch}$. For example, for long distances $G^{\bar{\rm a}}$ is expected to be low, and so it is easier for Eve to guarantee the condition $G \geq G^{\bar{\rm a}}$. However, one could compensate the low $G^{\bar{\rm a}}$ by reducing μ to make Alice's signals harder to identify, thereby reducing G. That is, $\mu_{\max}$ typically decreases with the channel distance. In particular, previous works [37, 43] showed that $\mu_{\max}$ scales linearly with $\eta_{\rm ch}$ when no technological constrains are placed on Eve. Since $\eta_{\rm ch}$ already appears in equation (36), the resulting scaling of $K_{\max}$ is quadratic with $\eta_{\rm ch}$.

To investigate if a technologically constrained eavesdropper can impose the same restrictive scaling, we plot $\mu_{\max}$ and $K_{\max}$ against $\eta_{\rm ch}$ in figure 7. We consider two scenarios, which we evaluate for two different values of ε each, and for both USD1 and USD2. In the first scenario, we assume that all of the devices used by Eve and the legitimate users are ideal, and the only imperfection is the mode mismatch characterized by ε. The second scenario considers the practical parameters introduced in table 1, and sets $\eta_B = \eta_E = 0.6$ (notice that this is a rather conservative assumption, since Eve's technological capabilities are expected to surpass those of the legitimate users). We set $f = 0.155$ and $t_B = 0.9$ in both instances. In terms of imperfect mode

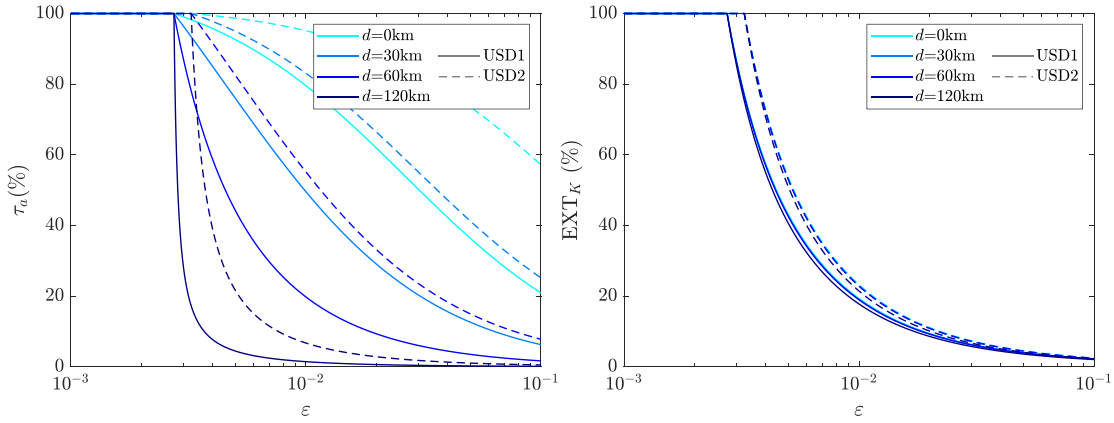

Figure 8. Fraction τ_a (in percentage) of the rounds that can be attacked by Eve while maintaining QBER ≤ 0.05 and $V_{\text{ave}} \geq 0.95$, and percentage of sifted key EXT_K that Eve can extract in such case, as a function of the error parameter ε and the transmission distance d between Alice and Bob. The parameter ε models Eve's imperfect mode overlap in her measurement implementation. For the simulations we set the attenuation coefficient of the channel to $\alpha_{\text{ch}} = 0.2\,\text{dB}\,\text{km}^{-1}$ and consider the parameters given in table 1, as well as the efficiency of Bob's detectors to be equal to that of Eve's. Note that the plots for the cases $d = 30$ km and $d = 60$ km in the right graph are almost overlapped for both USD1 and USD2, with the plot for $d = 0$ km just above them with some amount of overlap too.

overlap, we run the simulations for $\varepsilon = 0.002$, which is sufficiently low to guarantee that $V_{\text{ave}} > V^{\text{th}}$ for both USD1 and USD2 and typical values of α (see figure 6), and for $\varepsilon = 0.004$, which results in a V_{ave} slightly below V^{th}. The results are plotted for a range of η_{ch} that corresponds to channel lengths between 30 and 300 km when considering a typical fiber-loss coefficient $\alpha_{\text{ch}} = 0.2\,\text{dB}\,\text{km}^{-1}$.

In line with [37, 43], a quadratic scaling across the entire range of η_{ch} is observed for both the USD1 and USD2 when considering the ideal parameters and $\varepsilon = 0.002$. However, this is no longer the case when considering the parameters from table 1, for which the quadratic scaling is only observed for up to $\eta_{\text{ch}} \approx 10^{-2.5}$. When a smaller η_{ch} —i.e. longer channel—is considered, μ_{max} becomes so small that Eve's attack can no longer fulfill the condition $V_{\text{ave}} > V^{\text{th}}$. Therefore, as there is no need to reduce G to prevent Eve's attack from being undetectable, μ_{max} remains constant from there on, and the scaling of K_{max} turns from quadratic to linear over η_{ch}. Interestingly, even though the upper bound imposed by the USD1 is stricter than that imposed for the USD2, the latter remains in the quadratic regime for longer. This is because the performance of the USD2 is worse in terms of gain, but better in terms of visibility (see figures 5 and 6).

Moving to $\varepsilon = 0.004$, we see that the USD2 barely allows to maintain the upper bound in the quadratic regime, and indeed the scaling is linear for $\eta_{\text{ch}} \gtrapprox 10^{-1.5}$. Naturally, larger errors in the implementation imply that Eve requires a higher value of μ to successfully attack the system. Therefore, by keeping μ below this level, it is guaranteed that Eve's attack is detectable. Moreover, the implementation using USD1 is completely unable to impose a bound on the secret-key rate at any distance. This is because the metrics under attack improve with μ, but only up to a certain optimal point. This point may appear at a different value of μ for each metric, and is a consequence of several factors (e.g. very high values of μ may result in lots of erroneous clicks at Eve's detectors, preventing her attack from remaining undetectable). This means that if for one of these optimal intensities the metrics do not all fall within their acceptance regions, then the attack will be unsuccessful at any other intensity, and so $\mu_{\text{max}} \to \infty$ and $K_{\text{max}} \to \infty$.

It is clear that, according to our simple model, achieving a precise mode overlap within the beamsplitters is crucial to the success of Eve's attack. Nonetheless, an eavesdropper encountering challenges in this regard might still manage to extract some amount of information by launching a partial attack, in which she acts only on a reduced subset of the transmitted signals. Figure 8 illustrates, for different values of the channel distance d, the fraction τ_a of the rounds that can be attacked by Eve while still maintaining her success with respect to the QBER and V_{ave}, as well as the maximum percentage of sifted key that she can extract in this scenario (see appendix F for details about how these quantities are computed). In particular, this means that both metrics are maintained within their corresponding acceptance ranges. Here, we assume a typical fiber-based channel with attenuation coefficient $\alpha_{\text{ch}} = 0.2\,\text{dB}\,\text{km}^{-1}$. Figure 8 shows that a shorter channel length results in a higher amount of rounds that can be attacked by Eve. This is to be expected, since the metrics are calculated from all the rounds, those intercepted by Eve and those that are not. Consequently, the better measurement statistics observed by Bob at short distances from the unattacked rounds favorably impact the overall value of the metrics. Moreover, figure 8 also shows that the number of rounds that Eve may attack falls very sharply for longer channel lengths once the metrics of her attack are outside of the

acceptability ranges. Again, this is due to the degradation of the legitimate users' metrics, as it is clear that in the extreme case where the QBER and visibility obtained during the unattacked rounds are equal to the threshold values $\mathrm{QBER}^{\mathrm{th}}$ and V^{th}, no additional error is permissible, and so Eve cannot attack any rounds without becoming detectable unless her metrics are an improvement over those of Alice and Bob. Nevertheless, the amount of sifted-key bits distilled during the unattacked rounds also increases at short distances, meaning that the ratio of sifted key known to Eve decreases. For the scenario described by the parameters in table 1, these two opposite effects result in very little variation of EXT_K over the distance d, as can be seen at the right in figure 8. Crucially, the figure shows that Eve's attack has the potential to compromise the entire secret key at low values of ε. What is more, even for higher values of ε, Eve can still successfully attack an important fraction of the rounds and obtain part of the sifted key.

7. Conclusions

Security proofs of QKD typically consider that the eavesdropper's capabilities are only limited by the laws of quantum mechanics. Here, we have evaluated a less conservative scenario in which Eve is actually restricted by current technology. In particular, we have studied the feasibility of zero-error attacks against COW QKD in this framework. To do so, we have introduced two practical receivers to perform an USD measurement of Alice's emitted signals, which is the essential step of this type of attack.

Both proposed USD receivers are rather simple, and employ only linear passive optics, phase-space displacement operations and threshold single-photon detectors. We have derived analytical expressions for the expected values of the main metrics (i.e. the gain, the QBER and the visibilities) of a COW QKD protocol assuming realistic experimental conditions, i.e. as a function of the most relevant device imperfections of Eve's equipment. In doing so, we have found that the most critical experimental parameter seems to be the quality of interference between Alice's weak coherent pulses and Eve's strong light during her displacement operation. Overall, our results indicate that zero-error attacks could break the security of COW QKD with present-day technology, particularly if Alice and Bob only consider the observed average visibility in their monitoring line, as it is done e.g. in commercial setups and long-distance implementations of this scheme.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

This work was supported by Cisco Systems Inc., 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 the Grant No. PID2020-118178RB-C21, MICIN with funding from the European Union NextGenerationEU (PRTR-C17.I1) and the Galician Regional Government with own funding through the 'Planes Complementarios de I+D+I con las Comunidades Autónomas' in Quantum Communication, the European Union's Horizon Europe Framework Programme under the Marie Sklodowska-Curie Grant No. 101072637 (Project QSI), as listed by the UKRI though the Engineering and Physical Sciences Research Council (EPRSC) with Grant No. EP/X028313/1, and the project 'Quantum Security Networks Partnership' (QSNP, Grant Agreement No. 101114043). P v L acknowledges funding from the BMBF in Germany (QR.X and QuaPhySI) and from the EU/BMBF via QuantEra (ShoQC).

Appendix A. Model for the imperfect mode overlap

Let us consider a real beamsplitter with transmittance T in which the two input optical pulses are not perfectly mode-matched, and therefore do not interfere perfectly (see figure 9). This imperfect mode overlap can be simplistically modeled by splitting each of the two inputs into two modes [52]. One of these modes from each pulse interferes as desired, while the two remaining modes do not interfere and only have their amplitudes diminished by the beamsplitter. The fraction of light from the first (second) input that is perfectly matched and therefore interferes in the beamsplitter is determined by the parameter t_1 (t_2), and the total degree of overlap in the beamsplitter is defined as $\mathcal{M} := t_1 t_2$ [52].

In order to model the splitting of the first (second) input into two modes, namely one that is properly matched, say $\hat{a}_m$ ($\hat{b}_m$), and one that does not interfere, say $\hat{a}_p$ ($\hat{b}_q$), an ideal beamsplitter of transmittance t_1 (t_2) can be used, as illustrated in figure 9. After that, the modes that are properly matched (i.e. $\hat{a}_m$ and $\hat{b}_m$) interfere in an ideal beamsplitter with the original transmittance T, resulting in a new mode, say $\hat{m}$, at one

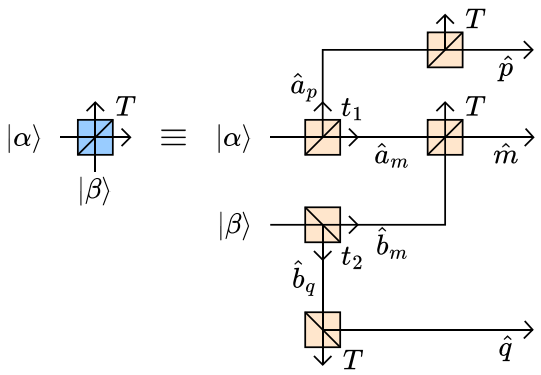

Figure 9. Simple model to characterize the imperfect mode overlap at one output of a beamsplitter with two coherent pulses $|\alpha\rangle$ and $|\beta\rangle$ as inputs. The original (blue) beamsplitter of transmittance T at the left, where pulses $|\alpha\rangle$ and $|\beta\rangle$ do not perfectly interfere, can be modeled as the network of perfectly mode-matched (orange) beamsplitters shown at the right. In this network, the input state $|\alpha\rangle$ ($|\beta\rangle$) is first split into two optical modes, $\hat{a}_m$ and $\hat{a}_p$ ($\hat{b}_m$ and $\hat{b}_q$), according to the parameter t_1 (t_2) that describes the corresponding beamsplitter transmittance. Subsequently, the optical modes $\hat{a}_m$ and $\hat{b}_m$ interfere in a beamsplitter with transmittance T, while each of the modes $\hat{a}_p$ and $\hat{b}_q$ interferes with vacuum in a beamsplitter with transmittance also T, resulting in the final output modes of the network $\hat{p}$, $\hat{m}$ and $\hat{q}$.

output of the beamsplitter. The other modes (i.e. $\hat{a}_p$ and $\hat{b}_q$) interfere with vacuum at ideal beamsplitters with transmittance T, resulting in new modes, say $\hat{p}$ and $\hat{q}$, at one output of each beamsplitter.

Appendix B. Derivation of the metrics

Section 5 of the main text provides all the necessary expressions to compute the expected values of the metrics for a system in the presence of Eve, given the configuration of the involved parties. It also includes a step-by-step derivation of the gain, G. This appendix provides the most relevant intermediate results for the derivation of both the QBER and visibilities, alongside with guidance on the interpretation of certain values.

B.1. QBER

Here we derive analytical expressions for the quantities $n_{\text{key}}(k)$ and $n_{\text{err}}(k)$ presented in equations (21) and (23). The derivation of equations (22) and (24) is straightforward.

First, we note that both $n_{\text{key}}(k)$ and $n_{\text{err}}(k)$ admit a decomposition similar to that used for $n_{\text{clk}}(k)$ in equation (14). That is, we can express

$$\begin{aligned} n_{\text{key}}(k) &= \sum_{i=0}^{2} p_{\mathbf{E}_i|\mathbf{E}_c}\, n_{\text{key}}(k|i)\,, \\ n_{\text{err}}(k) &= \sum_{i=0}^{2} p_{\mathbf{E}_i|\mathbf{E}_c}\, n_{\text{err}}(k|i)\,, \end{aligned} \tag{B1}$$

where $n_{\text{key}}(k|i)$ ($n_{\text{err}}(k|i)$) denotes the average number of key bits (erroneous key bits) distilled from a conclusive-block of length k given that the first signal in the block is identified by Eve as $|\varphi_i\rangle$.

Let $\mathbf{A}_{\text{key}}$ represent the event in which Alice emits a data signal, i.e. $\mathbf{A}_{\text{key}} = \mathbf{A}_0 \cup \mathbf{A}_1$. Then, it follows that $p_{\mathbf{A}_{\text{key}}} = p_{\mathbf{A}_0} + p_{\mathbf{A}_1}$, $p_{\mathbf{A}_{\text{key}}|\mathbf{E}_j} = p_{\mathbf{A}_0|\mathbf{E}_j} + p_{\mathbf{A}_1|\mathbf{E}_j}$, and $p_{\mathbf{A}_{\text{key}}|\mathbf{E}_c} = p_{\mathbf{A}_0|\mathbf{E}_c} + p_{\mathbf{A}_1|\mathbf{E}_c}$. Additionally, let $p_{\text{err}|\mathbf{B}_i}$ be the probability that Bob distills an incorrect key bit (i.e. a key bit different from the one sent by Alice), given that he receives $|\varphi_i\rangle$ from Eve. Note that this implies that the result of Eve's measurement for that signal was $\mathbf{E}_i$ and this signal was not discarded by her processing. Moreover, we define $p_{\text{err}|\mathbf{B}_c} = \sum_{i=0}^{2} p_{\text{err}|\mathbf{B}_i}$ as the probability that Bob distills an erroneous bit given that he received a signal different from $|\varphi_{\text{vac}}\rangle$. Then, we have that

$$\begin{aligned} p_{\text{err}|\mathbf{B}_0} &= p_{\mathbf{A}_0|\mathbf{E}_0}\frac{p_{\text{d}}^{\text{D}}}{2} + p_{\mathbf{A}_1|\mathbf{E}_0}\left(1 - \frac{p_{\text{d}}^{\text{D}}}{2}\right), \\ p_{\text{err}|\mathbf{B}_1} &= p_{\mathbf{A}_0|\mathbf{E}_1}\left(1 - \frac{p_{\text{d}}^{\text{D}}}{2}\right) + p_{\mathbf{A}_1|\mathbf{E}_1}\frac{p_{\text{d}}^{\text{D}}}{2}, \\ p_{\text{err}|\mathbf{B}_2} &= \frac{p_{\mathbf{A}_0|\mathbf{E}_2} + p_{\mathbf{A}_1|\mathbf{E}_2}}{2}. \end{aligned} \tag{B2}$$

With this, we can express both $n_{\text{key}}(k|i)$ and $n_{\text{err}}(k|i)$ in a recursive form. Specifically, we have that

$$\begin{aligned}
n_{\text{key}}(k|0) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\mathbf{A}_{\text{key}}|\mathbf{E}_0}\left(1+p_{\text{clk}|\mathbf{B}_v}\right)+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_c}(k-2)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left[\left(p_{\mathbf{A}_{\text{key}}|\mathbf{E}_0}+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_1}\right)p_{\text{clk}|\mathbf{B}_v}+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_c}(k-2)\right] \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\text{key}}(k-1|0)+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_2}p_{\text{clk}|\mathbf{B}_v}\right], \\
n_{\text{key}}(k|1) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\mathbf{A}_{\text{key}}|\mathbf{E}_0}+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_1}+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_c}(k-2)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left[p_{\mathbf{A}_{\text{key}}|\mathbf{E}_1}\left(1+p_{\text{clk}|\mathbf{B}_v}\right)+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_c}(k-2)\right] \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\text{key}}(k-1|1)+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_2}p_{\text{clk}|\mathbf{B}_v}\right], \\
n_{\text{key}}(k|2) = {} & n_{\text{key}}(k-1)+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_2}p_{\text{clk}|\mathbf{B}_v},
\end{aligned} \tag{B3}$$

and

$$\begin{aligned}
n_{\text{err}}(k|0) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\mathbf{A}_{\text{key}}|\mathbf{E}_0}p_{\text{err}|\mathbf{B}_v}+p_{\text{err}|\mathbf{B}_0}+p_{\text{err}|\mathbf{B}_c}(k-2)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left[\left(p_{\mathbf{A}_{\text{key}}|\mathbf{E}_0}+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_1}\right)p_{\text{err}|\mathbf{B}_v}+p_{\text{err}|\mathbf{B}_c}(k-2)\right] \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\text{err}}(k-1|0)+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_2}p_{\text{err}|\mathbf{B}_v}\right], \\
n_{\text{err}}(k|1) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\text{err}|\mathbf{B}_0}+p_{\text{err}|\mathbf{B}_1}+p_{\text{err}|\mathbf{B}_c}(k-2)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left[p_{\text{err}|\mathbf{B}_1}+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_1}p_{\text{err}|\mathbf{B}_v}+p_{\text{err}|\mathbf{B}_c}(k-2)\right] \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\text{err}}(k-1|1)+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_2}p_{\text{err}|\mathbf{B}_v}\right], \\
n_{\text{err}}(k|2) = {} & n_{\text{err}}(k-1)+p_{\mathbf{A}_{\text{key}}|\mathbf{E}_2}p_{\text{err}|\mathbf{B}_v}.
\end{aligned} \tag{B4}$$

The starting points for the recursions above are

$$\begin{aligned}
n_{\text{key}}(1|i) &= p_{\mathbf{A}_{\text{key}}|\mathbf{E}_i}p_{\text{clk}|\mathbf{B}_v}, \qquad n_{\text{key}}(0) = 0, \\
n_{\text{err}}(1|i) &= p_{\mathbf{A}_{\text{key}}|\mathbf{E}_i}p_{\text{err}|\mathbf{B}_v}, \qquad n_{\text{err}}(0) = 0,
\end{aligned} \tag{B5}$$

for $i \in \{0,1\}$. By solving the recursions, one finally obtains equations (21) and (23).

B.2. Visibilities

We focus first on the visibility V_2. According to equations (1) and (26), this visibility can be computed from the values of $n_{MX,2}(k)$ and $p^{\text{last}}_{MX,2}(k)$, with $X \in \{1,2\}$. It is straightforward to obtain the value of $p^{\text{last}}_{MX,2}(k)$, shown in equation (28), so we will focus here on $n_{MX,2}(k)$.

Let $p_{\mathbf{E}_i|\mathbf{A}_j,\mathbf{E}_c}$ denote the conditional probability that the outcome of Eve's USD measurement is $\mathbf{E}_i$, given that Alice prepared the signal $|\varphi_j\rangle$ and Eve's measurement was conclusive. That is,

$$p_{\mathbf{E}_i|\mathbf{A}_j,\mathbf{E}_c} = \frac{p_{\mathbf{E}_i|\mathbf{A}_j}}{p_{\mathbf{E}_0|\mathbf{A}_j}+p_{\mathbf{E}_1|\mathbf{A}_j}+p_{\mathbf{E}_2|\mathbf{A}_j}}, \tag{B6}$$

for $i,j \in \{0,1,2\}$. Now, we express $n_{MX,2}(k)$ in the form

$$n_{MX,2}(k) = \sum_{i=0}^{2} p_{\mathbf{E}_i|\mathbf{E}_c}\, n_{MX,2}(k|i), \tag{B7}$$

where $n_{MX,2}(k|i)$ is the average number of signals sent by Alice as $|\varphi_2\rangle$ that prompt a click in D_{MX} within a conclusive-block of length k, given that the first signal of the block is $|\varphi_i\rangle$. We can write recursive expressions to describe these quantities, such as

$$\begin{aligned}
n_{\text{M1},2}(k|0) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\mathbf{A}_2|\mathbf{E}_0}\left(1+p_{\text{d}}^{\text{M1}}\right)+p_{\mathbf{A}_2|\mathbf{E}_c}(k-2)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left[\left(p_{\mathbf{A}_2|\mathbf{E}_0}+p_{\mathbf{A}_2|\mathbf{E}_1}\right)p_{\text{d}}^{\text{M1}}+p_{\mathbf{A}_2|\mathbf{E}_c}(k-2)\right] \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\text{M1},2}(k-1|0)+p_{\mathbf{A}_2|\mathbf{E}_2}p_{\text{d}}^{\text{M1}}\right], \\
n_{\text{M1},2}(k|1) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\mathbf{A}_2|\mathbf{E}_0}+p_{\mathbf{A}_2|\mathbf{E}_1}+p_{\mathbf{A}_2|\mathbf{E}_c}(k-2)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left[p_{\mathbf{A}_2|\mathbf{E}_1}\left(1+p_{\text{d}}^{\text{M1}}\right)+p_{\mathbf{A}_2|\mathbf{E}_c}(k-2)\right] \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\text{M1},2}(k-1|1)+p_{\mathbf{A}_2|\mathbf{E}_2}p_{\text{d}}^{\text{M1}}\right], \\
n_{\text{M1},2}(k|2) = {} & n_{\text{M1},2}(k-1)+p_{\mathbf{A}_2|\mathbf{E}_2}p_{\text{d}}^{\text{M1}},
\end{aligned} \tag{B8}$$

and

$$\begin{aligned}
n_{\mathrm{M2},2}(k|0) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\mathbf{A}_2|\mathbf{E}_0}\left(1+p_{\mathrm{d}}^{\mathrm{M2}}\right)+p_{\mathbf{A}_2|\mathbf{E}_c}\left(1-p_{\mathbf{E}_2|\mathbf{A}_2,\mathbf{E}_c}\right)(k-2)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left[\left(p_{\mathbf{A}_2|\mathbf{E}_0}+p_{\mathbf{A}_2|\mathbf{E}_1}\right)p_{\mathrm{d}}^{\mathrm{M2}}+p_{\mathbf{A}_2|\mathbf{E}_c}\left(1-p_{\mathbf{E}_2|\mathbf{A}_2,\mathbf{E}_c}\right)(k-2)\right] \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\mathrm{M2},2}(k-1|0)+p_{\mathbf{A}_2|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M2}}\right], \\
n_{\mathrm{M2},2}(k|1) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\mathbf{A}_2|\mathbf{E}_0}+p_{\mathbf{A}_2|\mathbf{E}_1}+p_{\mathbf{A}_2|\mathbf{E}_c}\left(1-p_{\mathbf{E}_2|\mathbf{A}_2,\mathbf{E}_c}\right)(k-2)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left[p_{\mathbf{A}_2|\mathbf{E}_1}\left(1+p_{\mathrm{d}}^{\mathrm{M2}}\right)+p_{\mathbf{A}_2|\mathbf{E}_c}\left(1-p_{\mathbf{E}_2|\mathbf{A}_2,\mathbf{E}_c}\right)(k-2)\right] \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\mathrm{M2},2}(k-1|1)+p_{\mathbf{A}_2|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M2}}\right], \\
n_{\mathrm{M2},2}(k|2) = {} & n_{\mathrm{M2},2}(k-1)+p_{\mathbf{A}_2|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M2}},
\end{aligned} \tag{B9}$$

being the starting points of the recursions

$$n_{MX,2}(1|i) = p_{\mathbf{A}_2|\mathbf{E}_i}p_{\mathrm{d}}^{MX}, \qquad n_{MX,2}(0) = 0, \tag{B10}$$

for $X \in \{1,2\}$ and $i \in \{0,1\}$. By solving the recursions, one obtains equation (27).

Now we focus on the visibilities V_s of sequences $s = s_2s_1$ that contain two signals. From equation (29), we have that these visibilities can be computed from the quantities $n_{MX,s_2s_1}(k)$, $p^{\mathrm{last}}_{MX,s_2s_1}(k)$ and $p^{\mathrm{edge}}_{MX,s_2s_1}(k)$, introduced in section 5.3. The quantities $p^{\mathrm{last}}_{MX,s_2s_1}(k)$ and $p^{\mathrm{edge}}_{MX,s_2s_1}(k)$ are given in equation (33), and their derivation is straightforward. Thus we focus on the derivation of $n_{MX,s_2s_1}(k)$.

Following a similar approach as with the previous metrics, first we write

$$n_{MX,s_2s_1}(k) = \sum_{i=0}^{2} p_{\mathbf{E}_2|\mathbf{E}_c}\, n_{MX,s_2s_1}(k|i), \tag{B11}$$

and then we focus on finding the quantities $n_{MX,s_2s_1}(k|i)$. Now, let $p_{MX,s_2s_1|\mathbf{B}_c}$ denote the conditional probability that Alice originally prepares the sequence s_2s_1 and D_{MX} registers a click in the time slot associated with the interference between the last pulse of s_1 and the first pulse of s_2, given that Bob received a non-vacuum signal in both rounds. This implies that Eve measured both signals conclusively and they were not discarded during her processing. Then we have that

$$p_{\mathrm{M1},s_2s_1|\mathbf{B}_c} = p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}\left[1-p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M1}}\right)\right], \tag{B12}$$

and

$$p_{\mathrm{M2},s_2s_1|\mathbf{B}_c} = p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}\left[p_{\mathrm{d}}^{\mathrm{M2}}+\left(p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}+p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}-2p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}\right)\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]. \tag{B13}$$

Putting all together, we have that the recursive expressions for the required quantities, as well as the starting points of the recursions, are given by

$$\begin{aligned}
n_{\mathrm{M1},s_2s_1}(k|0) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left\{p_{\mathbf{A}_{s_1}|\mathbf{E}_0}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}\left[1-p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M1}}\right)\right]+p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_0}+p_{\mathrm{M1},s_2s_1|\mathbf{B}_c}(k-3)\right\} \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left\{p_{\mathbf{A}_{s_1}|\mathbf{E}_0}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}\left[1-p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M1}}\right)\right]\right. \\
& \left. + p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_1}\left[1-p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M1}}\right)\right]+p_{\mathrm{M1},s_2s_1|\mathbf{B}_c}(k-3)\right\} \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\mathrm{M1},s_2s_1}(k-1|0)+p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M1}}\right], \\
n_{\mathrm{M1},s_2s_1}(k|1) = {} & p_{\mathbf{E}_0|\mathbf{E}_c}\left[p_{\mathbf{A}_{s_1}|\mathbf{E}_1}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}+p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_0}+p_{\mathrm{M1},s_2s_1|\mathbf{B}_c}(k-3)\right] \\
& + p_{\mathbf{E}_1|\mathbf{E}_c}\left\{p_{\mathbf{A}_{s_1}|\mathbf{E}_1}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}+p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_1}\left[1-p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M1}}\right)\right]+p_{\mathrm{M1},s_2s_1|\mathbf{B}_c}(k-3)\right\} \\
& + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\mathrm{M1},s_2s_1}(k-1|1)+p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M1}}\right], \\
n_{\mathrm{M1},s_2s_1}(k|2) = {} & n_{\mathrm{M1},s_2s_1}(k-1)+p_{\mathbf{A}_{s_1}|\mathbf{E}_2}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}p_{\mathrm{d}}^{\mathrm{M1}},
\end{aligned} \tag{B14}$$

and

$$\begin{aligned}
n_{\mathrm{M2},s_2s_1}(k|0) &= p_{\mathbf{E}_0|\mathbf{E}_c}\Big\{p_{\mathbf{A}_{s_1}|\mathbf{E}_0}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}\left[1-p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]\\
&\quad + p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_0}\left[p_{\mathrm{d}}^{\mathrm{M2}}+p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]+p_{\mathrm{M2},s_2s_1|\mathbf{B}_c}(k-3)\Big\}\\
&\quad + p_{\mathbf{E}_1|\mathbf{E}_c}\Big\{p_{\mathbf{A}_{s_1}|\mathbf{E}_0}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}\left[1-p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]\\
&\quad + p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_1}\left[1-p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]+p_{\mathrm{M2},s_2s_1|\mathbf{B}_c}(k-3)\Big\}\\
&\quad + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\mathrm{M2},s_2s_1}(k-1|0)+p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M2}}\right],\\
n_{\mathrm{M2},s_2s_1}(k|1) &= p_{\mathbf{E}_0|\mathbf{E}_c}\Big\{p_{\mathbf{A}_{s_1}|\mathbf{E}_1}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}\left[p_{\mathrm{d}}^{\mathrm{M2}}+p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]\\
&\quad + p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_0}\left[p_{\mathrm{d}}^{\mathrm{M2}}+p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]+p_{\mathrm{M2},s_2s_1|\mathbf{B}_c}(k-3)\Big\}\\
&\quad + p_{\mathbf{E}_1|\mathbf{E}_c}\Big\{p_{\mathbf{A}_{s_1}|\mathbf{E}_1}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}\left[p_{\mathrm{d}}^{\mathrm{M2}}+p_{\mathbf{E}_1|\mathbf{A}_{s_2},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]\\
&\quad + p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_1}\left[1-p_{\mathbf{E}_0|\mathbf{A}_{s_1},\mathbf{E}_c}\left(1-p_{\mathrm{d}}^{\mathrm{M2}}\right)\right]+p_{\mathrm{M2},s_2s_1|\mathbf{B}_c}(k-3)\Big\}\\
&\quad + p_{\mathbf{E}_2|\mathbf{E}_c}\left[n_{\mathrm{M2},s_2s_1}(k-1|1)+p_{\mathbf{A}_{s_1}|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M2}}\right],\\
n_{\mathrm{M2},s_2s_1}(k|2) &= n_{\mathrm{M2},s_2s_1}(k-1)+p_{\mathbf{A}_{s_1}|\mathbf{E}_2}p_{\mathbf{A}_{s_2}|\mathbf{E}_c}p_{\mathrm{d}}^{\mathrm{M2}},
\end{aligned} \tag{B15}$$

and

$$\begin{aligned}
n_{\mathrm{M1},s_2s_1}(2|0) &= p_{\mathbf{A}_{s_1}|\mathbf{E}_0}\left[p_{\mathbf{E}_0|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_0}+p_{\mathbf{E}_1|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_1}p_{\mathrm{d}}^{\mathrm{M1}}+p_{\mathbf{E}_2|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M1}}\right],\\
n_{\mathrm{M1},s_2s_1}(2|1) &= p_{\mathbf{A}_{s_1}|\mathbf{E}_1}\left[p_{\mathbf{E}_0|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_0}+p_{\mathbf{E}_1|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_1}+p_{\mathbf{E}_2|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M1}}\right],\\
n_{\mathrm{M2},s_2s_1}(2|0) &= p_{\mathbf{A}_{s_1}|\mathbf{E}_0}\left[p_{\mathbf{E}_0|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_0}+p_{\mathbf{E}_1|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_1}p_{\mathrm{d}}^{\mathrm{M2}}+p_{\mathbf{E}_2|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M2}}\right],\\
n_{\mathrm{M2},s_2s_1}(2|1) &= p_{\mathbf{A}_{s_1}|\mathbf{E}_1}\left[p_{\mathbf{E}_0|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_0}p_{\mathrm{d}}^{\mathrm{M2}}+p_{\mathbf{E}_1|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_1}+p_{\mathbf{E}_2|\mathbf{E}_c}p_{\mathbf{A}_{s_2}|\mathbf{E}_2}p_{\mathrm{d}}^{\mathrm{M2}}\right],\\
n_{\mathrm{M1},s_2s_1}(0) &= n_{\mathrm{M2},s_2s_1}(0)=0.
\end{aligned} \tag{B16}$$

Appendix C. Alternative USD measurement

In this appendix we introduce an alternative USD setup for Eve that, unlike the USD1, allows her to identify the decoy signals $|\varphi_2\rangle$, but provides a lower overall success probability than USD1. We call this scheme USD2 and, as we did for the USD1, below we account for the most common imperfections in its model, and calculate its corresponding measurement statistics.

The idealized optical scheme is shown in figure 10(a). First, an optical displacement $\hat{D}(-\alpha/2)$ is applied to the pulses generated by Alice, transforming each pulse $|\chi\rangle$ into $|\chi-\frac{\alpha}{2}\rangle$. Subsequently, each displaced pulse enters a 50:50 beamsplitter, where it interferes with a coherent state $|\alpha/2\rangle$. The resulting state $|\chi/\sqrt{2}\rangle$ ($|(\chi-\alpha)/\sqrt{2}\rangle$) at the output port of the beamsplitter associated with constructive (destructive) interference is then detected by $D_{\mathrm{E}+}$ ($D_{\mathrm{E}-}$). Importantly, this means that, in the absence of imperfections, a click in $D_{\mathrm{E}+}$ ($D_{\mathrm{E}-}$) can only occur when Alice prepares a coherent (vacuum) pulse.

Since Alice's signals have two optical pulses, Eve has to implement this measurement twice, each time for a different pulse. The correspondence between each measurement result $\mathbf{E}_i$ and the pattern of detections needed to prompt it is shown in table 2.

The experimental setup for USD2 is shown in figure 10(b). As with USD1, in practice one can approximate the optical displacement $\hat{D}(-\alpha/2)$ with a beamsplitter of transmittance $T\approx 1$ together with an interfering offset coherent pulse $|\beta\rangle$ satisfying

$$\beta=-\sqrt{\frac{T}{1-T}}\frac{\alpha}{2}. \tag{C1}$$

With this choice of β, Eve's approximated displacement operation performs the transformation $|0\rangle\to|-\sqrt{T}\alpha/2\rangle$ and $|\alpha\rangle\to|\sqrt{T}\alpha/2\rangle$. Since this introduces some loss, we adjust the amplitude of the interfering signal $|\theta\rangle$ at the second beamsplitter to

$$\theta=\sqrt{T}\frac{\alpha}{2}. \tag{C2}$$

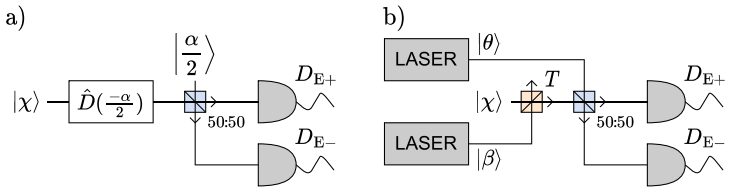

Figure 10. Alternative USD measurement setup for Eve, which we refer to as USD2. (a) Ideal optical model. (b) Proposed experimental realization. The values of θ and β are given in the text.

Table 2. Assignments between click patterns and measurement outcomes for USD2. $D_{E1\pm}$ ($D_{E2\pm}$) refers to the detector $D_{E\pm}$ acting in the first (second) optical pulse of a signal. Symbols '✓', 'X' and '–' mean that, for each measurement result, the indicated detector clicks, does not click, or is irrelevant, respectively.

	D_{E1+}	D_{E1-}	D_{E2+}	D_{E2-}
$\mathbf{E}_0$	—	X	X	✓
$\mathbf{E}_1$	X	✓	—	X
$\mathbf{E}_2$	✓	X	✓	X
$\mathbf{E}_3$	Otherwise			

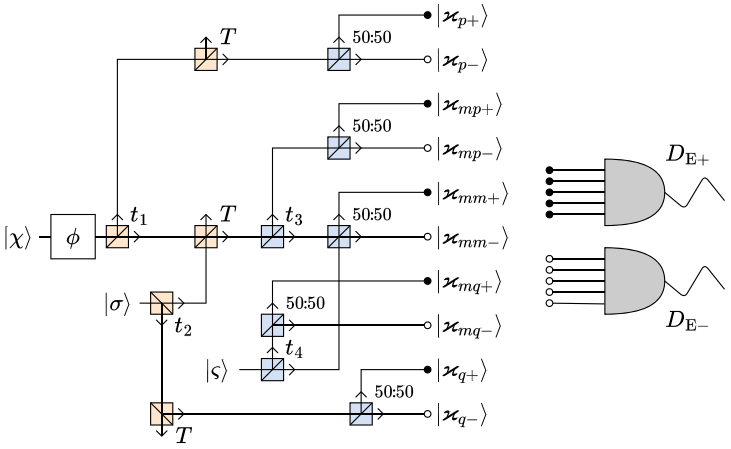

Figure 11. Model of the alternative USD measurement, USD2, including the main experimental imperfections. The beamsplitters depicted in orange (blue) are those related to the behavior of the first (second) beamsplitter of the experimental setup. Each output mode of the scheme marked with a black (white) circle is received by the detector D_{E+} (D_{E-}). That is, the total intensity received by each detector is the sum of the intensities in each of its corresponding input modes.

To fairly compare the performance of USD2 with that of USD1 in realistic scenarios, we consider the same imperfections in both setups. This includes the possibility of a small phase shift ϕ of the incoming pulses $|\chi\rangle$, the use of imperfect detectors with detection efficiency η_E and dark-count probability p_d^E, intensity fluctuations in the coherent pulses $|\beta\rangle$ and $|\theta\rangle$, and, lastly, a potentially imperfect mode overlap at each of the beamsplitters. To characterize this latter effect we use again the simple model from [52]. Moreover, for simplicity, we assume that the percentage of deviation in intensity is equal for $|\beta\rangle$ and $|\theta\rangle$, and modifies their values, respectively, by

$$\sigma = \sqrt{1+\delta}\beta, \qquad \varsigma = \sqrt{1+\delta}\theta, \tag{C3}$$

where $\delta \in [-1, \infty)$.

The complete model with imperfections is depicted in figure 11. Similar to the analysis of USD1, the parameter t_1 (t_2) denotes here the fraction of $|\chi\rangle$ ($|\sigma\rangle$) that correctly interferes at the first beamsplitter. Similarly, for the 50:50 beamsplitter, t_3 (t_4) represents the fraction of light in the first (second) input port of this beamsplitter that correctly interferes. For simplicity, we disregard any interference effect in the second beamsplitter between those optical modes that did not correctly interfere in the first one.

The total intensity received by D_{E+} (D_{E-}) given that Alice sends $|\chi\rangle$, namely $\mu_{E+|\chi}$ ($\mu_{E-|\chi}$), is the sum of the intensities at each independent optical mode within the constructive (destructive) output port of the second beamsplitter. That is, $\mu_{E\pm|\chi} = \sum_x |\varkappa_{x\pm}|^2$, where $|\varkappa_{x\pm}\rangle$ denotes the equivalent coherent state in the

output mode $x\pm$, with $x \in \{p, q, mp, mm, mq\}$ (see figure 11). Therefore, we have that

$$\mu_{\mathrm{E}\pm|\chi} = \frac{T}{4}\left[2|\chi|^2 + (1+\delta)\,|\alpha|^2\left(1 \mp \sqrt{t_2 t_3 t_4}\right) - 2\sqrt{t_1\,(1+\delta)}\left(\sqrt{t_2} \mp \sqrt{t_3 t_4}\right)\mathrm{Re}\left\{\chi\,\alpha^* e^{i\phi}\right\}\right]. \tag{C4}$$

Let $p_{\overline{\mathrm{E}\pm}|\chi=0}$ ($p_{\overline{\mathrm{E}\pm}|\chi=\alpha}$) be the probability that $D_{\mathrm{E}\pm}$ does not click given that Alice prepares $|0\rangle$ ($|\alpha\rangle$). This probabilities can be straightforwardly computed from $p_{\overline{\mathrm{E}\pm}|\chi} = (1 - p_{\mathrm{d}}^{\mathrm{E}})\exp\{-\eta_{\mathrm{E}}\mu_{\mathrm{E}\pm|\chi}\}$. Then, by particularizing equation (C4) to each possible input state, we obtain

$$\begin{aligned}
\mu_{\mathrm{E}\pm|\chi=0} &= \frac{T}{4}|\alpha|^2\,(1+\delta)\left(1 \mp \sqrt{t_2 t_3 t_4}\right), \\
\mu_{\mathrm{E}\pm|\chi=\alpha} &= \frac{T}{4}|\alpha|^2\left[2 + (1+\delta)\left(1 \mp \sqrt{t_2 t_3 t_4}\right) - 2\sqrt{t_1\,(1+\delta)}\left(\sqrt{t_2} \mp \sqrt{t_3 t_4}\right)\cos\phi\right].
\end{aligned} \tag{C5}$$

Finally, given the assignments presented in table 2, it is immediate to calculate the probabilities $p_{\mathbf{E}_i|\mathbf{A}_j}$ for the setup USD2 as

$$\begin{aligned}
p_{\mathbf{E}_0|\mathbf{A}_0} &= p_{\mathbf{E}_1|\mathbf{A}_1} = p_{\overline{\mathrm{E}-}|\chi=\alpha}\,p_{\overline{\mathrm{E}+}|\chi=0}\left(1 - p_{\overline{\mathrm{E}-}|\chi=0}\right), \\
p_{\mathbf{E}_0|\mathbf{A}_1} &= p_{\mathbf{E}_1|\mathbf{A}_0} = p_{\overline{\mathrm{E}-}|\chi=0}\,p_{\overline{\mathrm{E}+}|\chi=\alpha}\left(1 - p_{\overline{\mathrm{E}-}|\chi=\alpha}\right), \\
p_{\mathbf{E}_0|\mathbf{A}_2} &= p_{\mathbf{E}_1|\mathbf{A}_2} = p_{\overline{\mathrm{E}-}|\chi=\alpha}\,p_{\overline{\mathrm{E}+}|\chi=\alpha}\left(1 - p_{\overline{\mathrm{E}-}|\chi=\alpha}\right), \\
p_{\mathbf{E}_2|\mathbf{A}_0} &= p_{\mathbf{E}_2|\mathbf{A}_1} = p_{\overline{\mathrm{E}-}|\chi=\alpha}\,p_{\overline{\mathrm{E}-}|\chi=0}\left(1 - p_{\overline{\mathrm{E}+}|\chi=\alpha}\right)\left(1 - p_{\overline{\mathrm{E}+}|\chi=0}\right), \\
p_{\mathbf{E}_2|\mathbf{A}_2} &= p^2_{\overline{\mathrm{E}-}|\chi=\alpha}\left(1 - p_{\overline{\mathrm{E}+}|\chi=\alpha}\right)^2, \\
p_{\mathbf{E}_3|\mathbf{A}_i} &= 1 - \sum_{j=0}^{2} p_{\mathbf{E}_j|\mathbf{A}_i} \quad \text{for } i \in \{0,1,2\}.
\end{aligned} \tag{C6}$$

Appendix D. Effect of other imperfections besides mode mismatch

As mentioned in the main text, the crucial imperfection that determines the success of Eve's attack is the mode mismatch at her beamsplitters. To show this, we examine here how the remaining protocol and experimental parameters affect the expected value of the metrics in the presence of the attack. In particular, we fix all these parameters to the values used in the main text (see table 1) with the exception of the parameter we want to study in each particular case. Besides, as done in the main text, the parameters that quantify the quality of the mode overlap at Eve's beamsplitters are set to $t_1 = t_2 = t_3 = t_4 = 1 - \varepsilon$, where here we pick $\varepsilon = 2 \cdot 10^{-3}$. This value of ε sufficiently low to ensure that both USD1 and USD2 succeed if all the other parameters are fixed to the values shown in table 1. Once again, for simplicity, we consider the same dark-count probability p_{d} for all detectors, i.e. $p_{\mathrm{d}}^{\mathrm{E}} = p_{\mathrm{d}}^{\mathrm{D}} = p_{\mathrm{d}}^{\mathrm{M}X} = p_{\mathrm{d}}$, with $X \in \{1,2\}$.

D.1. Effect of the intensity of Alice's pulses

First, we investigate the influence of the intensity $\mu = |\alpha|^2$ of Alice's pulses on Eve's attack performance. This is shown in figure 12, where the expected values of the metrics as a function of the intensity μ are displayed. Not surprisingly, the gain rises in alignment with μ, as the probability of a conclusive measurement rises as well. The QBER and visibilities also improve as μ increases, as a higher gain reduces the impact of dark counts on Bob's detectors on these metrics. Notably, the QBER stays well below the critical threshold for all evaluated intensities.

Concerning the visibilities, we note that those computed for USD2 show a continuous growth with μ across the entire considered range. This is because this scheme can identify the three signals emitted by Alice, and the probability of a conclusive measurement increases with μ. However, for USD1, the visibility of sequences containing decoy signals only improves at low μ, stabilizing once the impact of the dark counts at Bob's detector becomes negligible. In particular, we note that the visibilities V_{02} and V_{21} reach this regime quicker than V_{22}, as the sequence '22' is more frequently resent as a vacuum signal.

D.2. Effect of the maximum length of the conclusive block, $M_{\max}$

Next we focus on the maximum length of a conclusive block, $M_{\max}$, which serves Eve to cap the memory resources required to record all the measurement results, as well as the time delay she has to introduce in the channel to apply her block processing strategy.

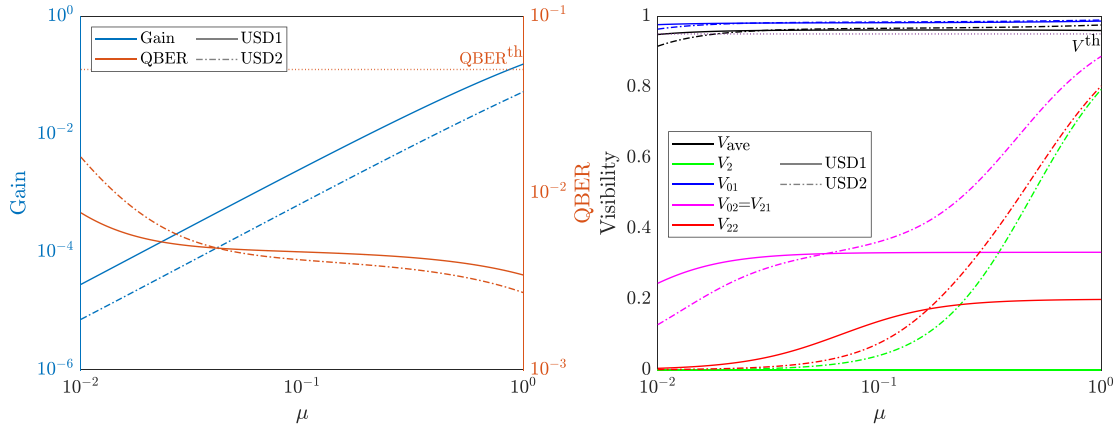

Figure 12. Gain, QBER and visibilities vs the intensity of Alice's pulses $\mu = |\alpha|^2$.

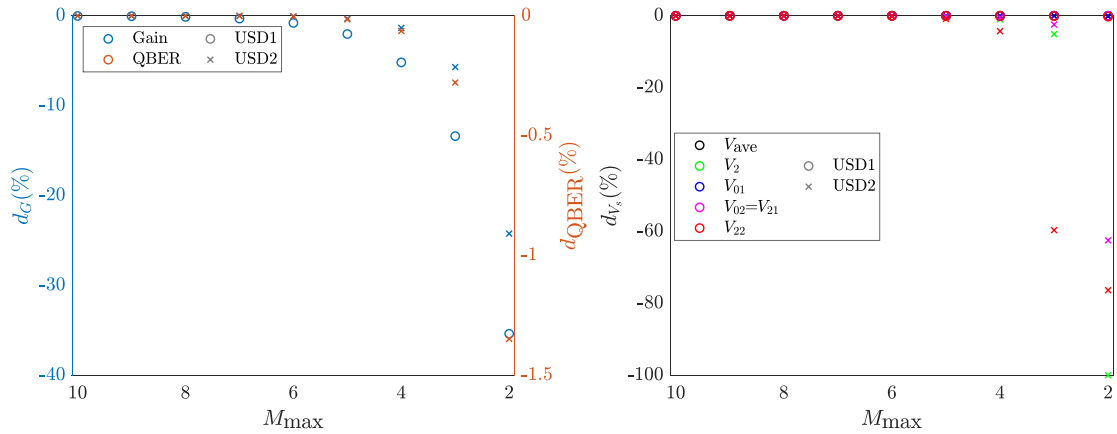

Figure 13. Relative difference of the gain (d_G), QBER (d_{QBER}) and visibilities (d_{V_s}), in percentage, for different values of $M_{\max}$ when compared to the same metrics with $M_{\max} \to \infty$, computed as in equation (D1), The intensity has been set to $\mu = 1$ and the case $M_{\max} \to \infty$ is approximated by $M_{\max} = 10^6$.

Previous analysis in [37, 43] asserted that the variation of the metrics with $M_{\max}$ is negligible for values of $M_{\max} > 10$ when considering practical values of α. To confirm this, we plot in figure 13 the relative difference $d_m(M_{\max})$ with respect to each metric $m \in \{G, \mathrm{QBER}, V_s\}$ for several values of $M_{\max}$, where

$$d_m(M_{\max}) := \frac{m(M_{\max}) - m(\infty)}{m(\infty)}, \tag{D1}$$

and $m(M_{\max})$ is defined as the value of the metric m if Eve's blocks are limited to $M_{\max}$ pulses, while $m(\infty)$ represents the value of that metric when there is no limit to the length of Eve's blocks (we approximate this by setting $M_{\max} = 10^6$).

The results shown in figure 13 are obtained considering $\mu = 1$ (that is, $|\alpha| = 1$), since this is close to the upper end of the practical values used in realistic COW implementations. It can be seen that the gain loses around a 35% of its value for USD1 when setting $M_{\max} = 2$, and around 25% for USD2. Nevertheless, the relative difference for values of $M_{\max} > 8$ is near zero, showing that $M_{\max} = 10$ indeed attains a similar performance as the asymptotic case, in terms of the gain.

On the other hand, the value of the QBER only varies up to $\approx$1% of its value when $M_{\max} = 2$, rapidly approaching zero difference when $M_{\max}$ increases, which proves that this metric does not degrade a lot when using short blocks. This is to be expected, as the actual probability of Eve introducing errors does not change for longer blocks, so the variation comes only from a lower significance of the effect of dark counts in Bob.

Lastly, we see that the change in the visibilities over $M_{\max}$ is quite small for USD1, and indeed the scale of this variation is similar to the one observed for the QBER. This is because this variation stems, once again, from an increase of dark counts in Bob when Eve sends only short blocks. However, the visibilities in USD2 do degrade significantly when the length of the blocks is more limited. In particular, the visibilities that deal with the decoy signal are the most affected, with V_2 degrading up to 100% when $M_{\max} = 2$. The reason is

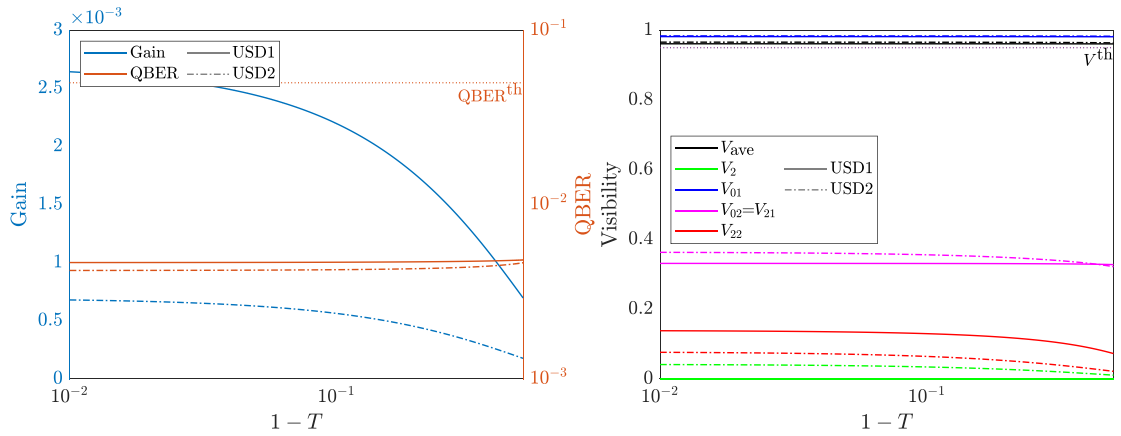

Figure 14. Gain, QBER and visibilities vs the transmittance of Eve's asymmetric beamsplitter, T.

that shorter blocks impose an artificial bias against the retransmission of decoy signals by Eve, even when properly identified, due to her processing strategy. In the limit of $M_{\max} = 2$, in fact, Eve never resends $|\varphi_2\rangle$, as this signal is translated into $|\varphi_{\text{vac}}\rangle$ when it is placed at the edge of a block. Thus, the decreased probability of resending a decoy signal makes the behavior of the USD2 similar to that observed in previous sections, where the blocks were short due to the small intensity, and therefore most of the signals sent to Bob when Alice sends $|\varphi_2\rangle$ correspond to Eve's misidentification of $|\varphi_2\rangle$ by one of the data signals, which naturally leads to poor visibility results.

In any case, it can be highlighted from figure 13 that all metrics are very close to their asymptotic value when setting $M_{\max} = 10$, even when Alice sends pulses with a relatively high intensity, and therefore it is sufficient for Eve to use this configuration in practical scenarios.

D.3. Effect of the transmittance of the asymmetric beamsplitter, T

Now we investigate the performance of Eve's attack as a function of the transmittance T of her asymmetric beamsplitter, which she uses to approximates an optical displacement in both USD measurement schemes. While this approximation is only accurate when $T \approx 1$, the simulations indicate that the value of T has a minimal impact on the attack's feasibility.

This is illustrated in figures 14(a) and (b), where the metrics are plotted against $1 - T$, within the range $(1 - T) \in [10^{-3}, 0.5]$. The simulations indicate that, as T decreases, so does the expected gain in both schemes, which is attributed to the small loss introduced by the approximate displacement. Importantly, however, the QBER and visibilities are largely unaffected by changes in T within a reasonable range, and both the QBER and V_{ave} remain acceptable even at $T = 0.5$. This stability is due to our specific choice of β.

Indeed, one could alternatively set $\beta = -\frac{\alpha}{\sqrt{1-T}}$ ($\beta = -\frac{\alpha}{2\sqrt{1-T}}$) for USD1 (USD2), and the resulting scheme would still approximate the desired displacement for $T \approx 1$. For instance, for USD1, this leads to the transformation $|0\rangle \to |-\alpha\rangle$ and $|\alpha\rangle \to |(\sqrt{T}-1)\alpha\rangle$. This choice avoids incurring into additional losses (thus resulting in a higher gain). However, this comes at the cost of increasing the probability of erroneous clicks in Eve's detector, making it a less convenient option for her. This is illustrated in figure 15, where we consider these alternative values of β, and call the resulting schemes as USD1* and USD2*.

D.4. Effect of the phase shift, ϕ

Figure 16 shows the dependency of the expected values of the metrics on the phase shift ϕ. Since the phase shift ϕ only has an impact when Alice transmits a coherent pulse, the decoy signal will be clearly influenced the most. In fact, the probability of misidentifying $|\varphi_2\rangle$ as a data signal increases with ϕ. As a consequence, signals that are never (in USD1) or rarely (in USD2) identified for low ϕ, result in conclusive measurements when this parameter increases, which in turn increases the gain. Of course, the probability of misidentifying a data signal also grows, which leads to an increment of the QBER, as shown in figure 16. Nevertheless, we note that the QBER remains relatively low even for phase shifts of several degrees.

Similarly, the visibilities generally decrease with ϕ, and in this case V_{ave} does fall below the acceptance threshold when $\phi \gtrapprox 2.5°$. Interestingly, V_{22} grows with ϕ until it reaches a certain point at which it stabilizes. This is again due to the increasing probability of misidentifying $|\varphi_2\rangle$ as one of the data signals. Note that, since the sequence '01' is favored by Eve's processing, the visibility of any two-signal sequence is expected to be nonzero even if the outcome of the USD measurements are totally random. In particular, in that extreme

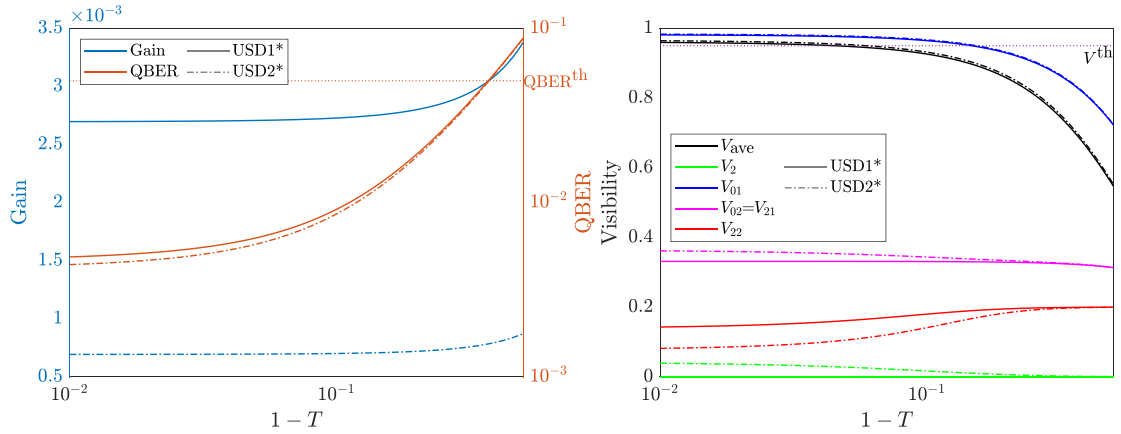

Figure 15. Gain, QBER and visibilities vs the transmittance of Eve's asymmetric beamsplitter, T, for the alternative schemes USD1* and USD2*. In particular, we consider that $\beta = -\frac{\alpha}{\sqrt{1-T}}$ ($\beta = -\frac{\alpha}{2\sqrt{1-T}}$) for USD1* (USD2*).

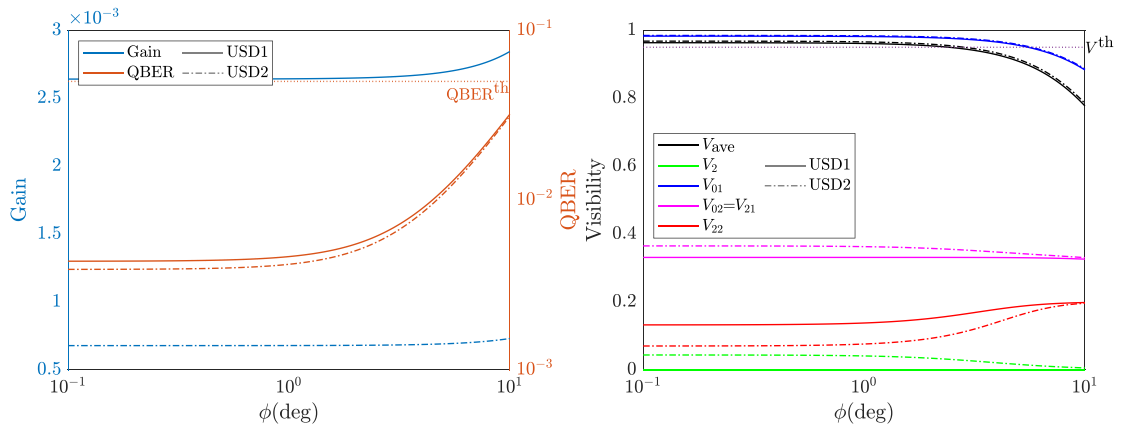

Figure 16. Gain, QBER and visibilities vs the phase shift of pulses received by Eve, ϕ (in degrees).

scenario, the sequence '22' could be identified as any other possible sequence by Eve, but the only two-signal block that she resends to Bob is '01', which always triggers the correct detector in the monitoring line, and hence increases V_{22}.

D.5. Effect of the intensity deviation, δ

Figure 17 illustrates the impact of small deviations in the intensities of Eve's pulses, quantified with the parameter δ, on the protocol metrics. Since δ can take both negative and positive values, as expressed in equations (4) and (C3), we plot it here in the range $[-0.3, 0.3]$, which corresponds to a deviation of $\pm 30\%$ over the intensity of the pulses.

Interestingly, the effects of positive or negative deviations are relatively distinct. In particular, the gain increases with δ through the entire depicted range. This is because higher intensities of $|\sigma\rangle$ and $|\varsigma\rangle$ result in larger click probabilities, especially in the case where Alice sends vacuum, since then Eve's signals are the only source of energy in the circuit. Notably, the QBER remains below its corresponding threshold even for significantly high deviations, although the results are worse for negative values of δ. This is because the probability of Eve observing a click when she is not supposed to (i.e. when a precise interference is intended to cancel the signal out) increases roughly as much for both positive and negative deviations. On the other hand, the same probability given that Eve is indeed supposed to observe a click grows with larger intensities, so the effect is relatively worse for negative deviations. Finally, figure 17 shows that the visibilities are more sensitive to δ. They exhibit a behavior similar to that shown in figures 6 and 16, with a sharper decrease in the value of the average visibility V_{ave} obtained for negative deviations, for similar reasons as for the QBER. In particular, V_{ave} is above the proposed acceptance threshold for deviations in the approximate range $\delta \in (-0.08, 0.1)$, which is still a relatively large margin.

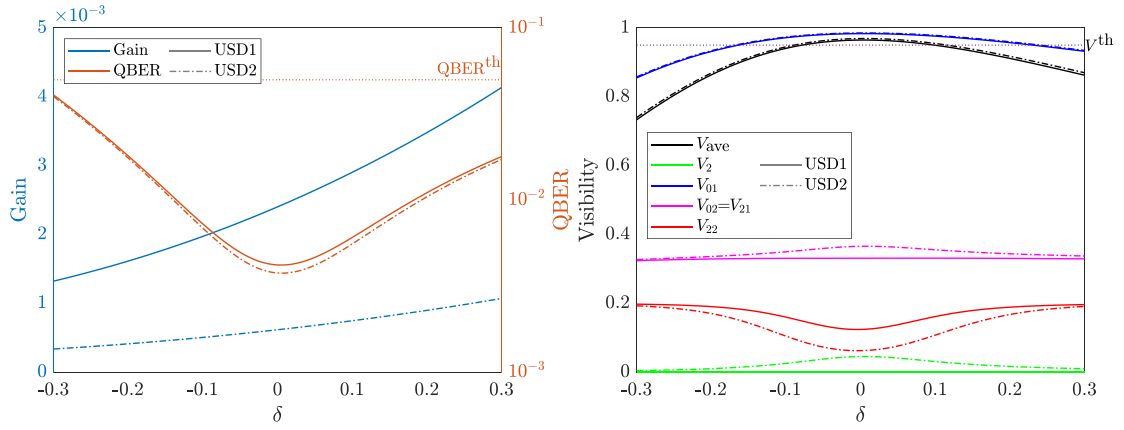

Figure 17. Gain, QBER and visibilities vs the intensity deviation of $|\beta\rangle$ (and $|\theta\rangle$ in USD2), δ.

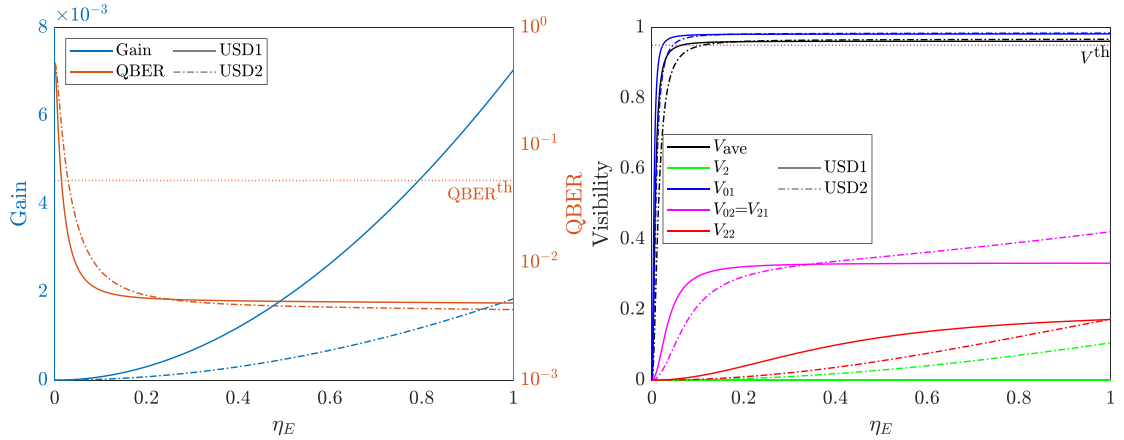

Figure 18. Gain, QBER and visibilities vs the efficiency of Eve's detectors, η_{E}.

D.6. Effect of the parameters of Eve's detectors

As expected, when η_E approaches zero, all metrics exhibit a degradation. This is because, in this scenario, Eve is sending a reduced number of signals to Bob, leading to a predominance of dark counts in Bob's system. This is illustrated in figure 18. The remaining conclusions drawn from this figure align with those inferred from figure 12. This is because the loss introduced by the nonideal detection efficiency of Eve's detector can be practically translated into an effective change in the intensity of Alice's signals to $\eta_{\mathrm{E}}\mu$.

The dependence of the metrics on the dark-count rate $p_{\mathrm{d}}^{\mathrm{E}}$ of Eve's detectors is shown in figure 19. The gain slightly increases for high values of $p_{\mathrm{d}}^{\mathrm{E}}$, as more erroneous clicks in the detectors lead to more conclusive measurements. Naturally, this comes with a degradation of the QBER and visibilities, although this degradation is quite gentle. In fact, the average visibility exhibits notable resilience to practical values of $p_{\mathrm{d}}^{\mathrm{E}}$, only falling below $V^{\mathrm{th}} = 0.95$ in the approximate range $p_{\mathrm{d}}^{\mathrm{E}} \gtrapprox 5 \cdot 10^{-5}$. As for the remaining visibilities, their behavior is relatively similar to the variation with ε, explained more in depth in appendix E, albeit much more mild in magnitude.

D.7. Effect of the parameters of Bob's detectors

Regarding Bob's detectors, we disregard the effect of their detection efficiency, as Eve can always make them click by resending him pulses with sufficiently high intensity. Thus, we focus on the impact of the dark counts.

Figure 20(a) shows the variation of the gain and QBER with the dark-count probability $p_{\mathrm{d}}^{\mathrm{D}}$ at D_{D}. Naturally, the gain rises with $p_{\mathrm{d}}^{\mathrm{D}}$, as it directly increases the number of clicks at Bob's data line. Moreover, the QBER also grows, as more random clicks lead to more errors. Unsurprisingly, this dependence is stronger than that observed with $p_{\mathrm{d}}^{\mathrm{E}}$, as the dark counts at Bob's detectors more directly cause errors than those at Eve's detectors. Indeed, for our particular choice of experimental and protocol parameters, an attack with USD2 remains viable only for $p_{\mathrm{d}}^{\mathrm{D}} \lessapprox 4 \cdot 10^{-5}$, while USD1 can sustain an attack for any plotted value of $p_{\mathrm{d}}^{\mathrm{D}}$.

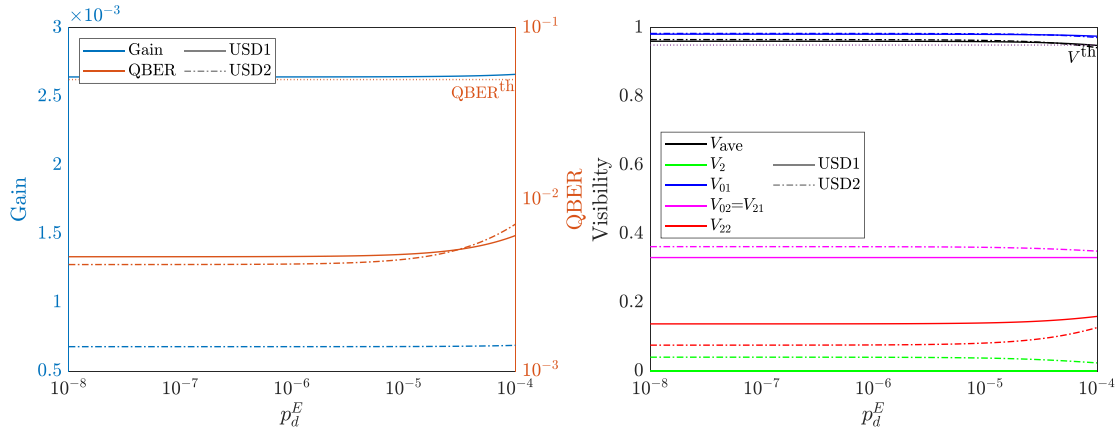

Figure 19. Gain, QBER and visibilities vs the dark-count probability of Eve's detectors, $p_{\mathrm{d}}^{\mathrm{E}}$. For simplicity, here we assume the same dark-count probability for all of them.

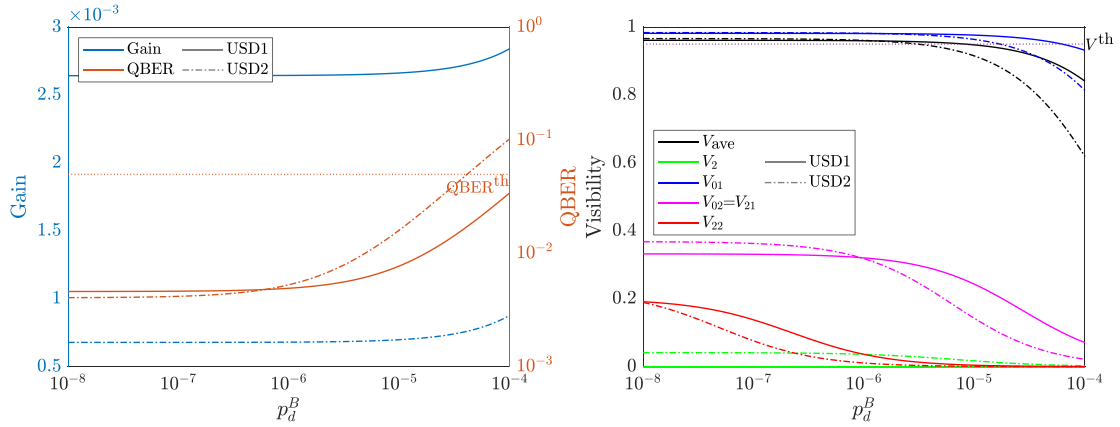

Figure 20. Gain, QBER and visibilities vs the probability of dark counts in Bob's corresponding detectors ($p_{\mathrm{d}}^{\mathrm{D}} = p_{\mathrm{d}}^{\mathrm{B}}$ for the gain and QBER, $p_{\mathrm{d}}^{\mathrm{M1}} = p_{\mathrm{d}}^{\mathrm{M2}} = p_{\mathrm{d}}^{\mathrm{B}}$ for the visibilities).

Regarding the visibilities, we consider for simplicity that both detectors at the monitoring line are equal, and so $p_{\mathrm{d}}^{\mathrm{M1}} = p_{\mathrm{d}}^{\mathrm{M2}} = p_{\mathrm{d}}^{MX}$. As expected, just like the QBER, all of them get worse as p_{d}^{MX} increases. Still, for the parameters we consider here, the average visibility remains sufficiently high given that $p_{\mathrm{d}}^{MX} \lessapprox 3 \cdot 10^{-6}$.

D.8. Effect of the probability of preparing the decoy signal, *f*

Finally, we investigate the effect of the decoy probability *f* on the metrics. As shown in figure 21, the gain decreases to nearly zero for large values of *f*. This is not only because Eve finds it more challenging to conclusively measure the decoy signal, but also and more importantly, because Eve's processing needs data signals to be located at the edges of the blocks. Thus, a small number of data signals sent by Alice means that most conclusive measurements come from decoy signals, which cannot be resent on their own. On the other hand, a smaller p_{E_c} also leads to a greater effect of clicks due to dark counts in Bob over the metrics, and indeed the QBER slightly grows for large values of *f*.

Regarding the visibilities, figure 21 showcases how the average visibility falls towards 0 for increasing values of *f*. This is because the probability of sending sequences that involve a decoy increases, which makes visibilities observing these sequences to have more weight over the final result of V_{ave}. On the other hand, the values of the individual visibilities do not decrease as significantly, and in fact, they remain essentially constant for the case of USD1. For USD2, however, those visibilities that depend on Alice sending a decoy signal slightly decrease with *f*. The reason for this is somewhat counter-intuitive. As more decoy signals are sent, blocks processed by Eve become smaller, so relatively more of the decoy signals that are not turned to vacuum come from erroneously measuring them as data signals, thus leading to a decreased visibility.

In any case, selecting a high value of *f* also severely decreases the secret-key rate of the protocol, as less data signals are emitted. Indeed, typical experiments of COW-QKD use a value of *f* quite low, around 0.15.

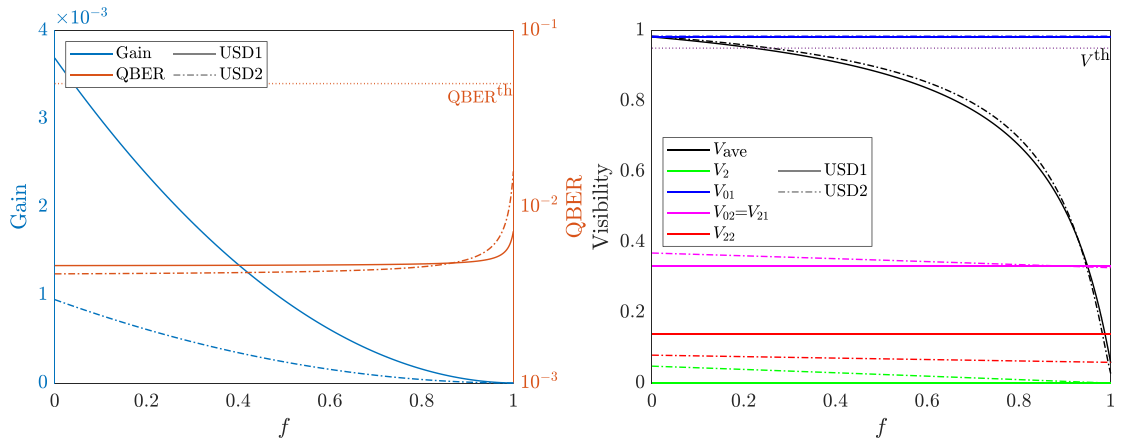

Figure 21. Gain, QBER and visibilities vs the transmission frequency of decoy signals, f.

Appendix E. Behavior of the visibilities V_{22}, V_{21}, and V_{02}

When ε is very small, most instances of $|\varphi_2\rangle$ (all of them if we consider the strategy USD1) are resent to Bob as vacuum signals, so the dark counts of Bob's detectors bring the visibilities down to zero.

As ε increases, so does the probability of conclusively (albeit erroneously) identifying a decoy signal $|\varphi_2\rangle$ as $|\varphi_0\rangle$ or $|\varphi_1\rangle$. In this scenario, Eve sends Bob more non-vacuum pulses, and relatively less clicks in the monitoring line are attributed to the dark counts. In the regime where dark counts are, comparatively to signal clicks, very low, the vacuum pulses sent by Eve—outside and at the edges of conclusive-blocks—can be ignored. Since USD1 never returns $\mathbf{E}_2$, and USD2 has a significantly higher probability of misidentifying $|\varphi_2\rangle$ as a data signal than correctly identifying $|\varphi_2\rangle$ for sufficiently high ε, we can analyze the visibilities, in this regime, by focusing on the four possible sequences of data signals that Eve can erroneously identify: '00', '01', '10' and '11'. In particular, when Bob receives '01', only D_{M1} can click due to the interference of two coherent pulses. The sequence '10' contains two vacuum pulses in the intermediate time slots, so no clicks can be observed at Bob (aside from those from dark counts). Finally, both '00' and '11' interfere $|0\rangle$ and $|\gamma\rangle$, resulting in a click in both D_{M1} and D_{M2} with very high probability.

Let us focus now in V_{22}. We notice that, whenever Alice sends '22', Eve misidentifies this sequence as one of the four previous sequences with equal probability, due again to the symmetry of the setup. This means that D_{M1} clicks with probability 3/4, while D_{M2} clicks with probability 2/4 (including possible double clicks in both detectors). Therefore, by using the definition of the visibility, we find that $V_{22} \approx 0.2$ for large values of ε.

Similar arguments apply to V_{02}. When Alice sends the sequence '02', we can distinguish between two different scenarios. If ε is high enough such that the dark counts are not the main source of clicks in the monitoring line, but the probability of mistaking one data signal for another is still sufficiently low, then Eve mostly misidentifies the original sequence as '00' or '01' with equal probability, and therefore $V_{02} \approx 1/3$. When ε increases, misidentifications of the data signal $|\varphi_0\rangle$ of the sequence '02' happen more often, and thus also the sequences '10' and '11', which implies that the visibility decreases. Analogous reasoning applies to V_{21}.

In addition, from figure 6 we observe that the range of values of ε where dark counts are relevant is larger for USD2 than for USD1, due to the fact that USD1 offers a larger $p_{\mathbf{E}_c}$, and therefore more non-vacuum pulses are resent. Moreover, this range is also considerably larger for V_{22} than for V_{02}, since two consecutive decoy signals are less likely to yield a conclusive measurement outcome than a single one. We also notice that USD2 performs better than USD1 for large values of ε. The reason for this behavior is that the probability of erroneously identifying $|\varphi_0\rangle$ as $|\varphi_1\rangle$ (or viceversa), given that the measurement is conclusive, is slightly smaller in USD2.

Appendix F. Partial attack

As shown in section 6, the average visibility V_{ave} in the presence of Eve's attack is above the acceptance threshold only for rather low values of the parameter ε. Nevertheless, it is still possible for Eve to remain undetected while obtaining partial information about the secret key. To this end, she can perform her attack on only a fraction of the rounds, so that the statistics from the unattacked rounds enhance the expected

values of the protocol metrics, compensating for the errors introduced by her attack. Here we explain how we evaluate the expected value of the metrics when Eve executes the attack on a fraction τ_a of the rounds.

We assume that the rounds under attack are consistently clustered in large groups of consecutive rounds, allowing us to disregard any possible border effects between the unattacked and attacked signals. Then we have that the gain is simply given by $G = \tau_a G^{\mathrm{a}} + (1-\tau_a) G^{\bar{\mathrm{a}}}$, where the superscript 'a' indicates that it is calculated for the system that is being attacked all the protocol rounds, as described in section 5, while '$\bar{\mathrm{a}}$' indicates that the metric is calculated in the absence of Eve. The result of $G^{\bar{\mathrm{a}}}$ can be computed from equation (35). To calculate the QBER, on the other hand, we have to find the values of N_{key} and N_{err}. Precisely, we can express these as $N_{\mathrm{key}} = \tau_a N^{\mathrm{a}}_{\mathrm{key}} + (1-\tau_a) N_{\mathrm{sig}} p^{\bar{\mathrm{a}}}_{\mathrm{key}}$ and $N_{\mathrm{err}} = \tau_a N^{\mathrm{a}}_{\mathrm{err}} + (1-\tau_a) N_{\mathrm{sig}} p^{\bar{\mathrm{a}}}_{\mathrm{err}}$, where

$$\begin{aligned} p^{\bar{\mathrm{a}}}_{\mathrm{key}} &= (1-f)\left[1-\left(1-p^{\mathrm{D}}_{\mathrm{d}}\right)^2 e^{-\eta_B \eta_{\mathrm{ch}} t_B |\alpha|^2}\right], \\ p^{\bar{\mathrm{a}}}_{\mathrm{err}} &= (1-f)\left[1+\left(1-p^{\mathrm{D}}_{\mathrm{d}}\right) e^{-\eta_B \eta_{\mathrm{ch}} t_B |\alpha|^2}\right] \frac{p^{\mathrm{D}}_{\mathrm{d}}}{2}, \end{aligned} \tag{F1}$$

are the probabilities of these events in the absence of Eve. Similarly, we can modify the values of the visibilities by making $N_{MX,s} = \tau_a N^{\mathrm{a}}_{MX,s} + (1-\tau_a) N_{\mathrm{sig}} p^{\bar{\mathrm{a}}}_{MX,s}$, where

$$\begin{aligned} \frac{p^{\bar{\mathrm{a}}}_{\mathrm{M1},2}}{f} &= \frac{p^{\bar{\mathrm{a}}}_{\mathrm{M1},s_2 s_1}}{p_{\mathbf{A}_{s_1}} p_{\mathbf{A}_{s_2}}} = 1-\left(1-p^{\mathrm{M1}}_{\mathrm{d}}\right) e^{-2\eta_B \eta_{\mathrm{ch}} (1-t_B) |\alpha|^2}, \\ \frac{p^{\bar{\mathrm{a}}}_{\mathrm{M2},2}}{f} &= \frac{p^{\bar{\mathrm{a}}}_{\mathrm{M2},s_2 s_1}}{p_{\mathbf{A}_{s_1}} p_{\mathbf{A}_{s_2}}} = p^{\mathrm{M2}}_{\mathrm{d}}, \end{aligned} \tag{F2}$$

are the probabilities corresponding to the relevant clicks when Eve does not act on the channel.

In order to compute the ratio of sifted key, EXT_K, that Eve can extract by enabling her attack during a fraction τ_a of all the communication rounds, one can observe that

$$\mathrm{EXT}_K = \frac{\tau_a N^{\mathrm{a}}_{\mathrm{key}}}{\tau_a N^{\mathrm{a}}_{\mathrm{key}} + (1-\tau_a) N_{\mathrm{sig}} p^{\bar{\mathrm{a}}}_{\mathrm{key}}}. \tag{F3}$$

ORCID iDs

Javier Rey-Domínguez https://orcid.org/0009-0006-2210-5470
Álvaro Navarrete https://orcid.org/0000-0003-3506-6037
Peter van Loock https://orcid.org/0000-0001-9445-0771
Marcos Curty https://orcid.org/0000-0002-0330-6828

References

[1] Lo H-K, Curty M and Tamaki K 2014 *Nat. Photon.* **8** 595
[2] Xu F, Ma X, Zhang Q, Lo H-K and Pan J-W 2020 *Rev. Mod. Phys.* **92** 025002
[3] Pirandola S *et al* 2020 *Adv. Opt. Photon.* **12** 1012
[4] Stucki D *et al* 2011 *New J. Phys.* **13** 123001
[5] Sasaki M *et al* 2011 *Opt. Express* **19** 10387
[6] Qiu J 2014 *Nature* **508** 441
[7] Chen Y-A *et al* 2021 *Nature* **589** 214
[8] Takeoka M, Guha S and Wilde M M 2014 *Nat. Commun.* **5** 5235
[9] Pirandola S, Laurenza R, Ottaviani C and Banchi L 2017 *Nat. Commun.* **8** 15043
[10] Jain N, Stiller B, Khan I, Elser D, Marquardt C and Leuchs G 2016 *Contemp. Phys.* **57** 366
[11] Marquardt C *et al* 2023 Implementation attacks against QKD systems *Technical Report* 575 Bundesamt für Sicherheit in der Informationstechnik, 53133 Bonn
[12] Hwang W-Y 2003 *Phys. Rev. Lett.* **91** 057901
[13] Lo H-K, Ma X and Chen K 2005 *Phys. Rev. Lett.* **94** 230504
[14] Wang X-B 2005 *Phys. Rev. Lett.* **94** 230503
[15] Lo H-K, Curty M and Qi B 2012 *Phys. Rev. Lett.* **108** 130503
[16] Comandar L C, Lucamarini M, Fröhlich B, Dynes J F, Sharpe A W, Tam S W-B, Yuan Z L, Penty R V and Shields A J 2016 *Nat. Photon.* **10** 312
[17] Woodward R I, Lo Y S, Pittaluga M, Minder M, Paraiso T K, Lucamarini M, Yuan Z L and Shields A J 2021 *npj Quantum Inf.* **7** 58
[18] Cao Y *et al* 2020 *Phys. Rev. Lett.* **125** 260503
[19] Wei K *et al* 2020 *Phys. Rev.* X **10** 031030
[20] Yin H-L *et al* 2016 *Phys. Rev. Lett.* **117** 190501
[21] Lucamarini M, Yuan Z L, Dynes J F and Shields A J 2018 *Nature* **557** 400
[22] Curty M, Azuma K and Lo H-K 2019 *npj Quantum Inf.* **5** 64
[23] Ma X, Zeng P and Zhou H 2018 *Phys. Rev.* X **8** 031043
[24] Lin J and Lütkenhaus N 2018 *Phys. Rev.* A **98** 042332

[25] Wang X-B, Yu Z-W and Hu X-L 2018 *Phys. Rev.* A **98** 062323
[26] Gisin N, Ribordy G, Zbinden H, Stucki D, Brunner N and Scarani V 2004 arXiv:quant-ph/0411022
[27] Stucki D, Brunner N, Gisin N, Scarani V and Zbinden H 2005 *Appl. Phys. Lett.* **87** 194108
[28] Stucki D, Walenta N, Vannel F, Thew R T, Gisin N, Zbinden H, Gray S, Towery C R and Ten S 2009 *New J. Phys.* **11** 075003
[29] Korzh B, Lim C C W, Houlmann R, Gisin N, Li M J, Nolan D, Sanguinetti B, Thew R and Zbinden H 2015 *Nat. Photon.* **9** 163
[30] Huttner B, Imoto N, Gisin N and Mor T 1995 *Phys. Rev.* A **51** 1863
[31] Brassard G, Lütkenhaus N, Mor T and Sanders B C 2000 *Phys. Rev. Lett.* **85** 1330
[32] ID Quantique SA, Geneva, Switzerland (available at: http://www.idquantique.com)
[33] Moroder T, Curty M, Lim C C W, Thinh L P, Zbinden H and Gisin N 2012 *Phys. Rev. Lett.* **109** 260501
[34] Lavie E and Lim C C-W 2022 *Phys. Rev. Appl.* **18** 064053
[35] Gao R-Q, Xie Y-M, Gu J, Liu W-B, Weng C-X, Li B-H, Yin H-L and Chen Z-B 2022 *Opt. Express* **30** 23783
[36] Li M-Y, Cao X-Y, Xie Y-M, Yin H-L and Chen Z-B 2024 *Phys. Rev. Res.* **6** 013022
[37] González-Payo J, Trényi R, Wang W and Curty M 2020 *Phys. Rev. Lett.* **125** 260510
[38] Waks E, Takesue H and Yamamoto Y 2006 *Phys. Rev.* A **73** 012344
[39] Curty M, Zhang L L, Lo H K and Lütkenhaus N 2007 *Quantum Inf. Comput.* **7** 665
[40] Tsurumaru T 2007 *Phys. Rev.* A **75** 062319
[41] Curty M, Tamaki K and Moroder T 2008 *Phys. Rev.* A **77** 052321
[42] Curty M, Lewenstein M and Lütkenhaus N 2004 *Phys. Rev. Lett.* **92** 217903
[43] Trényi R and Curty M 2021 *New J. Phys.* **23** 093005
[44] Chefles A 1998 *Phys. Lett.* A **239** 339
[45] Chefles A and Barnett S M 1998 *Phys. Lett.* A **250** 223
[46] Branciard C, Gisin N, Lütkenhaus N and Scarani V 2007 *Quantum Inf. Comput.* **7** 639
[47] Wehner S, Schaffner C and Terhal B M 2008 *Phys. Rev. Lett.* **100** 220502
[48] König R, Wehner S and Wullschleger J 2012 *IEEE Trans. Inf. Theory* **58** 1962
[49] Damgård I B, Fehr S, Renner R, Salvail L and Schaffner C 2007 A Tight High-Order Entropic Quantum Uncertainty Relation with Applications *Advances in Cryptology - CRYPTO 2007* (*Berlin, Heidelberg*) ed A Menezes (Springer) pp 360–78
[50] Wehner S, Curty M, Schaffner C and Lo H-K 2010 *Phys. Rev.* A **81** 052336
[51] Paris M G 1996 *Phys. Lett.* A **217** 78
[52] Laiho K, Avenhaus M, Cassemiro K N and Silberhorn C 2009 *New J. Phys.* **11** 043012