Two-way Gaussian quantum cryptography against coherent attacks in direct reconciliation

Carlo Ottaviani,[1,*] Stefano Mancini,[2] and Stefano Pirandola[1]

[1]*Department of Computer Science and York Center for Quantum Technologies, University of York, York YO10 5GH, United Kingdom*

[2]*School of Science and Technology, University of Camerino, 62032 Camerino, Italy, and INFN Sezione di Perugia, 06123 Perugia, Italy*

(Received 26 August 2015; published 11 December 2015)

We consider a two-way quantum cryptographic protocol with coherent states assuming direct reconciliation. A detailed security analysis is performed considering a two-mode coherent attack, which represents the residual eavesdropping once the parties have reduced the general attack by applying symmetric random permutations. In this context we provide a general analytical expression for the key rate, discussing the impact of the residual two-mode correlations on the security of the scheme. In particular, we identify the optimal eavesdropping against two-way quantum communication, which is given by a two-mode coherent attack with symmetric and separable correlations.

DOI: 10.1103/PhysRevA.92.062323 PACS number(s): 03.67.Dd, 03.65.−w, 42.50.−p, 89.70.Cf

I. INTRODUCTION

The goal of quantum key distribution (QKD) [1] is to make available unconditionally secure private keys between two authenticated users, Alice and Bob. Carriers of the information are quantum systems whose quantum nature is exploited to generate the same random sequence of bits, then to be used as a cryptographic key in one-time pad protocols. This strategy is based on the fundamental restriction, imposed by quantum mechanics, that obtaining perfect copies of arbitrary quantum states is impossible. In fact, any attempt in this sense unavoidably introduces some noise perturbing the quantum state itself (no-cloning theorem [2]).

To convert this feature of the quantum world into the ultimate cipher [3], any quantum cryptographic protocol needs to be arranged in a first quantum communication step, followed by a classical communication one. During the first stage, Alice encodes classical information into nonorthogonal quantum states, which are sent to Bob over a noisy quantum channel. This is used N times and assumed to be in the hands of an eavesdropper (Eve). The quantum signals are measured by Bob, detecting a noisy version of Alice's quantum states. After many uses of the channel ($N \gg 1$), the parties can share a random sequence of bits called the raw key. At this point, the parties sacrifice part of the N bits, from the raw key, communicating over a classical public channel. This allows them to compare the data in their hands and to estimate the presence of the eavesdropper on the quantum channel. This second stage allows Alice and Bob to quantify the adequate amount of error correction and privacy amplifications needed to reduce the stolen information to a negligible amount [4].

In recent years, continuous variable (CV) quantum systems [5,6] have attracted increasing attention for the implementation of quantum communication tasks, with special attention devoted to Gaussian CV states. The appealing possibilities of this approach are based on the replacement of single-photon pulses with bright coherent states and single-photon detection with simpler and more efficient Gaussian operations like homodyne and/or heterodyne detection schemes. This simplifies the experimental realization, on one hand, and can increase the key-rate production of the protocols by many orders of magnitude, on the other [1,7–10]. Furthermore, Gaussian CV protocols can easily go broadband. Within this research area, quantum cryptography has been one of the most prolific fields in recent years [6], with extensive theoretical and experimental research developed to improve the performances of point-to-point communications in one-way [11,12] and two-way [13,14] protocols.

In two-way schemes the parties exploit twice the quantum channel per each use of the protocol [13,14] (see also Ref. [15] for Discrete Variable (DV) two-way protocols and Refs. [16–18] for CV two-way protocols based on quantum illuminations [19–21]). In particular, CV two-way protocols [13,14] can achieve higher security thresholds thanks to an improved tolerance to the eavesdropper's noise. In fact, the analysis developed in Ref. [13] (see, for example, Fig. 3 in Ref. [13]) proved that, for fixed values of the channel's transmissivity, CV two-way protocols tolerate a higher level of noise than one-way protocols in the case of collective attacks. This makes this approach appealing to achieve high-rate secure communication in noisier environments, where one-way communication fails to provide a secure key.

In this work we study the security of two-way QKD considering general coherent attacks and focusing on direct reconciliation. In this case (see Fig. 1) Gaussian-modulated reference coherent states, $|\beta\rangle$, are sent from Bob to Alice through the quantum channel and are processed by Alice via a random displacement operation, $D(\alpha)$, with Gaussian modulation of amplitudes α. The output $\rho(\alpha,\beta)$ is sent backward to Bob, who applies heterodyne detection and classical postprocessing, in order to subtract the reference amplitude β and infer Alice's signal amplitude α. The higher tolerance to noise, granted by the double-use of the quantum channel, is due to the fact that Eve needs to attack both the forward and the backward steps of the quantum communication, in order to extract information on both β and α [13].

The key rate of the two-way QKD protocol has been studied under the standard assumption of collective Gaussian attacks [6]. Protection against coherent attacks can be achieved by switching randomly between single and double-use of the quantum channel (on-off switching) [13]. Collective attacks mean that Eve attaches uncorrelated ancillary modes to each

*Carlo.Ottaviani@york.ac.uk

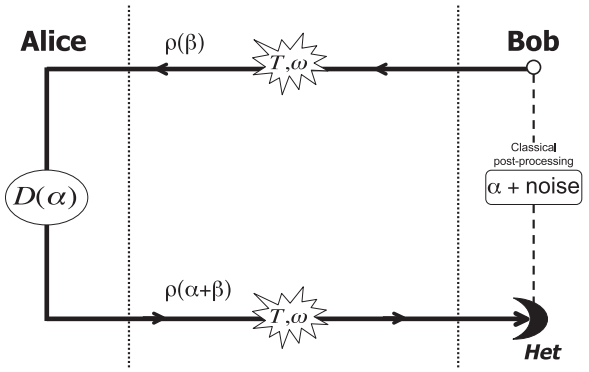

FIG. 1. In the general two-way protocol Bob sends the reference state $\rho(\beta)$ to Alice, who applies a random displacement $D(\alpha)$. The resulting Gaussian state $\rho(\alpha+\beta)$ is sent back to Bob, who applies heterodyne detection and classical postprocessing to recover Alice's encoding (α).

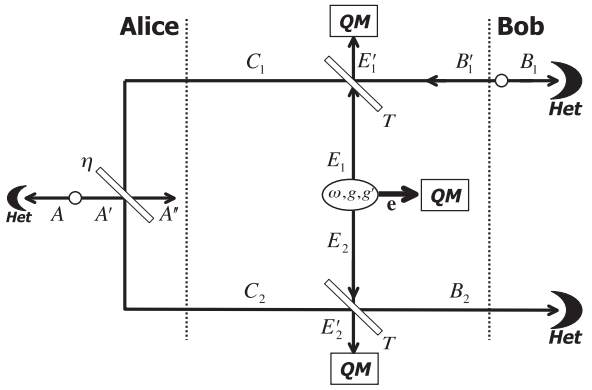

FIG. 2. Entanglement-based representation of the two-way QKD protocol. Bob prepares reference coherent states $|\beta\rangle$. This can be done by heterodyning one part of an EPR state. One mode is measured (B_1), while the other, B_1', is sent to Alice through an insecure quantum channel. Alice applies a random displacement of the reference state, $D(\alpha)$, which can be implemented by a beam splitter with transmissivity η and another EPR state. Choosing appropriately the transmissivity η and the variance of her EPR state, Alice sends displaced output state $\rho(\alpha,\beta)$ back to Bob. These are heterodyned and classically postprocessed by Bob. In this way he recovers Alice's encoding by subtracting the known reference amplitude β. The information encoded in the amplitude α is then used to obtain the raw key.

use of the quantum channel. The ancillae interact unitarily with the communication modes and are then measured by the eavesdropper. In this scenario, recently, it has been possible to extend two-way QKD also to the case where the parties encode information affected by trusted thermal noise [14].

In the present study we explicitly derive the secret key of the two-way protocol in the case where Eve's ancillary states are correlated. Inthis case the Alice-Bob communication line becomes a memory channel [22,23], in contrast to the case of collective attacks, where it is memoryless. Here we report a security analysis of a two-way CV-QKD protocol against coherent attacks. Our analysis is based on the conventional assumption that the parties exchange a large number of signals ($N \gg 1$). In this case we can reduce the general attack to a simpler two-mode coherent attack where, for each use of the protocol, Eve's ancillae share nonzero two-mode correlations. In addition to this, we consider the case of asymptotically large Gaussian modulation of the amplitudes α and β. This allows us to work with analytical mathematical expressions and to find the optimal two-mode coherent attack against the protocol, when Eve injects symmetric separable correlations [22].

The results for the two-way protocol are compared with the performances of the one-way version of the scheme and show that eavesdropping two-way quantum communication with a suitable two-mode coherent attack can reduce the performances partly below the one-way security threshold. This represents an example of a coherent attack overcoming the performances of collective ones, in point-to-point protocols. We discuss why this happens, in the context considered here, and finally, we compare our results with those of other recent studies [7,8,24] where two-mode optimal coherent attacks have been identified for end-to-end cryptographic protocols.

Our results are important for the development of the security analysis of CV protocols and for identification of the general challenges to implementing secure point-to-point communications. Our results confirm that the on-off switching operated by Alice, described in detail in Refs. [13] and [24], represents a necessary countermeasure to overcome the problem of realistic coherent attacks in two-way point-to-point quantum cryptography.

The structure of this paper is as follows. In Sec. II we introduce the protocol and illustrate the reduction of the general eavesdropping to a two-mode coherent attack. In Sec. III we provide the definition of the key rate and we show how to compute the Holevo bound and Alice-Bob mutual information, arriving at the analytical expression of the secret-key rate. In Sec. IV we analyze the security thresholds and we study the behavior of the relevant quantities as a function of Eve's injected thermal noise and degree of two-mode correlation. Section V presents our conclusions.

II. PROTOCOL AND EAVESDROPPING

We show the protocol in the entanglement-based representation (see Fig. 2). We reduce the general coherent eavesdropping to two-mode coherent attacks, and we illustrate the steps to compute the total and conditional covariance matrices. Then in the next section we provide the analytical expression of the symplectic spectra, which are used to compute the Holevo bound.

A. Coherent Gaussian attack

In a general (coherent) eavesdropping, Eve processes all N uses of the quantum channel applying a global coherent unitary operation that correlates all the modes involved in the different uses. However, exploiting the quantum de Finetti theorem [25] for infinite-dimensional systems, this general scenario can be reduced to a two-mode coherent attack. The parties can apply symmetric random permutations of the classical data in such a way that for $N \gg 1$, the cross correlations between distinct uses of the two-way communication can be neglected. The global coherence of the attack is thus reduced to a two-mode

coherence, between the forward and the backward channels involved in each round-trip quantum communication.

This residual two-mode coherent attack, in the most typical case, is implemented by two beam splitters of transmissivity T [26], where Eve mixes two ancillary modes, E_1 and E_2 (see Fig. 2). These two ancillae belong to a generally larger set of modes, $\{E_1, E_2, \mathbf{e}\}$, defining the pure initial quantum state owned by the eavesdropper. The two-mode Gaussian state $\rho_{E_1E_2}$ is generally correlated and described by the following covariance matrix (CM):

$$\mathbf{V}_{E_1E_2} = \begin{pmatrix} \omega\mathbf{I} & \mathbf{G} \\ \mathbf{G} & \omega\mathbf{I} \end{pmatrix}, \quad \mathbf{G} := \begin{pmatrix} g & 0 \\ 0 & g' \end{pmatrix}. \tag{1}$$

Here the parameter ω describes the variance of the thermal noise injected by Eve in the beam splitters, $\mathbf{I} = \mathrm{diag}(1,1)$, $\mathbf{Z} = \mathrm{diag}(1,-1)$, and matrix $\mathbf{G}$ accounts for the specific two-mode correlations employed by Eve to eavesdrop. The parameters ω, g, and g' must fulfill the conditions given in Ref. [22], in order to represent a physical attack. Note that the properties of this type of non-Markovian channel have recently been studied in the context of relay-based CV quantum cryptography [7,8], where they have also been classified and grouped into three possible cases. More recently it has been shown how they could be exploited to reactivate entanglement distribution and quantum communication protocols [27].

We distinguish three possible extremal cases: *collective attacks*, for $g = g' = 0$, corresponding to the standard collective eavesdropping; *separable attacks*, defined by the condition $|g| = |g'| = \omega - 1$, representing coherent attacks with separable correlations injected; and finally, *Einstein-Podolski-Rosen (EPR) attacks*, where $g = -g' = \sqrt{\omega^2 - 1}$ and $g = -g' = -\sqrt{\omega^2 - 1}$. These three eavesdropping strategies are not equivalent, and in the next section we identify the optimal one.

B. Entanglement-based protocol

We perform the security analysis in the entanglement-based representation so that, besides previous dilation of the quantum channel, we also provide the purification of the source of Bob's coherent states and Alice's random displacements. Thus, by referring to Fig. 2, we first assume that Bob's coherent states originate from two-mode squeezed vacuum states (EPR states), which are zero-mean Gaussian state, is described by the CM

$$\mathbf{V}_{B_1B_1'} = \begin{pmatrix} \mu_B\mathbf{I} & \sqrt{\mu_B^2 - 1}\mathbf{Z} \\ \sqrt{\mu_B^2 - 1}\mathbf{Z} & \mu_B\mathbf{I} \end{pmatrix}, \tag{2}$$

where the variance parameter μ_B quantifies the entanglement and also the local thermal noise in modes B_1 and B_1'. The heterodyne measurement performed by Bob on mode B_1 remotely projects mode B_1' on a coherent state traveling forward (from Bob to Alice) through the quantum channel. Its amplitude is classically modulated with a Gaussian distribution having variance $\mu = \mu_B - 1$.

At Alice's station the random displacement $D(\alpha)$ can be implemented by means of a beam splitter of transmissivity η. This mixes the incoming mode C_1 with a mode A', coming from Alice's EPR pairs A and A', whose Gaussian quantum state, $\rho_{AA'}$, is described by the following CM:

$$\mathbf{V}_{AA'} = \begin{pmatrix} \mu_A\mathbf{I} & \sqrt{\mu_A^2 - 1}\mathbf{Z} \\ \sqrt{\mu_A^2 - 1}\mathbf{Z} & \mu_A\mathbf{I} \end{pmatrix}. \tag{3}$$

While Alice's mode A' is sent through the beam splitter, the other mode A is heterodyne detected, in order to project the mode A' onto a coherent state $|\gamma\rangle$ modulated with variance μ_γ such that

$$\mu_\gamma = \mu_A - 1. \tag{4}$$

This setup is a way to equivalently simulate Alice's random displacements. In fact, for simplicity, consider the case where Eve is absent and there is no loss or noise in the quantum channel. In this scenario Alice receives $|\beta\rangle$ and must send $|\beta + \alpha\rangle = D(\alpha)|\beta\rangle$ back to Bob. We can see that, using the setup with the beam splitter, Alice prepares her output mode C_2 in the coherent state

$$|\sqrt{\eta}\beta + \sqrt{1-\eta}\gamma\rangle. \tag{5}$$

Now, in order to obtain a coherent state of the form $|\beta + \alpha\rangle$ from Eq. (5), we design Alice's beam splitter to have transmissivity $\eta \to 1$, and we assume that the coherent amplitude $\gamma \to \infty$ in such a way that

$$\gamma = \frac{\alpha}{\sqrt{1-\eta}}.$$

This is possible in theory by using an EPR input state for Alice with divergent variance $\mu_\gamma + 1$, where

$$\mu_\gamma := \frac{\mu}{1-\eta}. \tag{6}$$

Under these assumptions we get

$$|\sqrt{\eta}\beta + \sqrt{1-\eta}\gamma\rangle \simeq |\beta + \alpha\rangle.$$

III. KEY RATE, HOLEVO FUNCTION, AND MUTUAL INFORMATION

In direct reconciliation the parties use Alice's amplitudes α to prepare the secret key. This means that, during the classical procedure of parameter estimation, error correction, and privacy amplification, Bob infers the values of Alice's variables α from the results of his measurements. The security performances are quantified by the asymptotic secret-key rate

$$R := I_{AB} - \chi_{EA}, \tag{7}$$

which is defined as the difference between Alice-Bob's mutual information I_{AB} and the Holevo function χ_{EA}, which upper bounds Eve-Alice's mutual information.

The advantage of using the entanglement-based representation in Sec. II B relies on the fact that we do not need to know the details of the coherent operations performed by Eve on the modes. Instead, we can compute the function χ_{EA} from the output quantum state of Alice and Bob [6]. More precisely, we compute Eve's Holevo information as

$$\chi_{EA} = S_E - S_{E|\alpha}, \tag{8}$$

where S_E is the von Neumann entropy of Eve's total output modes, which coincides with the von Neumann entropy of Alice's and Bob's total output modes B_1, A, A'', and B_2. The other quantity is the von Neumann entropy of Eve's output modes conditioned on Alice's detection α. This is equal to the von Neumann entropy of Bob's output modes B_1 and B_2 conditioned on α.

For Gaussian states, the von Neumann entropy has a particularly simple form in terms of the symplectic eigenvalues [6]. It is given by

$$S := \sum_{\nu} h(\nu), \tag{9}$$

where ν are the symplectic eigenvalues of the CM associated with the state, and the entropic function $h(\nu)$ is defined as

$$h(\nu) := \frac{\nu+1}{2}\log_2\frac{\nu+1}{2} - \frac{\nu-1}{2}\log_2\frac{\nu-1}{2}.$$

This expression simplifies further in the limit of large modulation $\mu \gg 1$, in which case we have

$$h(\nu) \to \log_2\frac{e}{2}\nu + O(\nu^{-1}). \tag{10}$$

In the next subsection we provide the total and conditional CMs corresponding to $\rho_{B_1AA''B_2}$ and $\rho_{B_1B_2|\alpha}$ and the respective symplectic spectra, which are then used to compute the Holevo bound χ_{EA}.

A. Total symplectic spectrum

The global Alice-Bob quantum state, $\rho_{B_1AA''B_2}$, is a Gaussian state whose properties are described by the CM (we use the modes ordering $B_1AA''B_2$)

$$\mathbf{V} = \begin{pmatrix} \mu_B\mathbf{I} & & \phi\mathbf{Z} & \theta\mathbf{Z} \\ & \mu_A\mathbf{I} & \xi\mathbf{Z} & \tau\mathbf{Z} \\ \phi\mathbf{Z} & \xi\mathbf{Z} & k\mathbf{I} & \delta\mathbf{I} \\ \theta\mathbf{Z} & \tau\mathbf{Z} & \delta\mathbf{I} & \varepsilon\mathbf{I} \end{pmatrix} + \begin{pmatrix} & & & \\ & & & \\ & & g_\delta\mathbf{G} & \\ & & & g_\varepsilon\mathbf{G} \end{pmatrix}, \tag{11}$$

where the missing matrix entries are 0 and we have defined

$$\begin{aligned} \phi &:= -\sqrt{T(1-\eta)\left(\mu_B^2-1\right)}, \\ \theta &:= T\sqrt{\eta\left(\mu_B^2-1\right)}, \\ k &:= \eta\mu_A + (1-\eta)[T\mu_B + (1-T)\omega], \\ \xi &:= \sqrt{\eta\left(\mu_A^2-1\right)}, \\ \tau &:= \sqrt{T(1-\eta)\left(\mu_A^2-1\right)}, \\ \varepsilon &:= T^2\eta\mu_B + T(1-\eta)\mu_A + (T\eta+1)(1-T)\omega, \\ g_\varepsilon &:= 2(1-T)\sqrt{\eta T}, \\ \delta &:= \sqrt{T\eta(1-\eta)}[\mu_A - T\mu_B - (1-T)\omega], \\ g_\delta &:= -(1-T)\sqrt{(1-\eta)}. \end{aligned} \tag{12}$$

To obtain the symplectic spectrum of the CM of Eq. (11), we first compute the matrix

$$\mathbf{M}_T = i\Omega\mathbf{V}, \tag{13}$$

where $\Omega = \bigoplus_{k=1}^{4}\tilde{\omega}_k$, with $\tilde{\omega}_k = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ the symplectic form. Then we compute the standard eigenvalues of Eq. (13). After performing simple algebra and taking the limit of large modulation ($\mu \gg 1$), we find the general expressions

$$\nu_1 = \sqrt{(\omega-g)(\omega-g')}, \tag{14}$$

$$\nu_2 = \sqrt{(\omega+g)(\omega+g')}, \tag{15}$$

$$\nu_3\nu_4 = (1-T)^2\mu^2, \tag{16}$$

where the dependency on the correlation parameter, g and g', generalizes the known total symplectic spectrum under collective attacks [13], recovered for $g = g' = 0$. Using this spectrum with Eqs. (9) and (10), one easily obtains the asymptotic total von Neumann entropy, which we can write as

$$S_E = h(\nu_1) + h(\nu_2) + \log_2\frac{e^2}{4}(1-T)^2\mu^2. \tag{17}$$

B. Conditional symplectic spectrum and Holevo bound

When the protocol is used in direct reconciliation Bob's conditional CM can be obtained straightforwardly considering the CM involving Bob's modes, obtained from Eq. (11) tracing out Alice's modes. This approach considerably simplifies the problem. Starting from the matrix

$$\mathbf{V}_{B_1B_2} = \begin{pmatrix} \mu_B\mathbf{I} & \theta\mathbf{Z} \\ \theta\mathbf{Z} & \varepsilon\mathbf{I} + g_\varepsilon\mathbf{G} \end{pmatrix}, \tag{18}$$

we set $\mu_A = 1$ to simulate the conditioning on Alice's measurements, to arrive at the conditional CM given by

$$\mathbf{V}_C = \mathbf{V}_{B_1B_2}(\mu_A = 1). \tag{19}$$

From this CM we compute the matrix

$$\mathbf{M}_C = i\Omega\mathbf{V}_C, \tag{20}$$

where $\Omega = \bigoplus_{k=1}^{2}\tilde{\omega}_k$, and we derive its spectrum. Considering the asymptotic limit for large μ and the limit $\eta \to 1$, we obtain the following pair of symplectic eigenvalues:

$$\begin{aligned} \bar{\nu}_1 &= \sqrt{\omega + 2g\frac{\sqrt{T}}{1+T}}\sqrt{\omega + 2g'\frac{\sqrt{T}}{1+T}}, \\ \bar{\nu}_2 &= (1-T^2)\mu. \end{aligned} \tag{21}$$

Using $\bar{\nu}_1$ and $\bar{\nu}_2$ in Eq. (9) and (10), we derive the conditional von Neumann entropy:

$$\begin{aligned} S_{E|\alpha} &= h(\bar{\nu}_1) + h(\bar{\nu}_2), \\ &= h(\bar{\nu}_1) + \log_2\frac{e}{2}(1-T^2)\mu. \end{aligned} \tag{22}$$

Finally, putting together the results of Eqs. (17) and (22) in the definition of the Holevo function, Eq. (8), we find the analytic expression of the Holevo bound:

$$\chi_{EA} = h(\nu_1) + h(\nu_2) - h(\bar{\nu}_1) + \log_2\frac{e}{2}\frac{1-T}{1+T}\mu. \tag{23}$$

C. Mutual information

To obtain the secret-key rate we also need Alice-Bob mutual information. Since both quadratures, q and p, of mode B_2 are measured, the mutual information I_{AB} is given by the expression

$$I_{AB} = \frac{1}{2}\log_2 \frac{V_B^q + 1}{V_{B|\alpha\beta}^q + 1} + \frac{1}{2}\log_2 \frac{V_B^p + 1}{V_{B|\alpha\beta}^p + 1},$$

where V_B^q and V_B^p represent the variances for quadratures q and p of mode B_2, while $V_{B|\alpha\beta}^q$ and $V_{B|\alpha\beta}^p$ describe the conditional variances after Bob and Alice's measurements. The former can be obtained from the diagonal block of the CM given in Eq. (18), describing mode B_2. This is given by the expression

$$\mathbf{B}_2 = \varepsilon \mathbf{I} + g_\varepsilon \mathbf{G}, \tag{24}$$

from which, taking the limit $\eta \to 1$ and setting $\mu_B = 1$, we obtain

$$V_B^q = T^2 + T\mu + (1 - T^2)\omega + 2g(1 - T)\sqrt{T},$$
$$V_B^p = T^2 + T\mu + (1 - T^2)\omega + 2g'(1 - T)\sqrt{T}.$$

The conditional variances, $V_{B|\alpha\beta}^q$ and $V_{B|\alpha\beta}^p$, can now be obtained setting $\mu = 0$ in the previous equations. Taking the limit of large modulation $\mu \gg 1$, we get the asymptotic Alice-Bob mutual information

$$I_{AB} = \frac{1}{2}\log_2 \frac{T^2\mu^2}{\sigma\sigma'}, \tag{25}$$

where

$$\sigma := V_{B|\alpha\beta}^q + 1 = \Delta + 2g(1 - T)\sqrt{T},$$
$$\sigma' := V_{B|\alpha\beta}^p + 1 = \Delta + 2g'(1 - T)\sqrt{T},$$

and

$$\Delta := 1 + T^2 + (1 - T^2)\omega.$$

D. Secret-key rate

We now have all the quantities needed to compute the secret-key rate defined in Eq. (7). From the expressions for the asymptotic mutual information given in Eq. (25) and the Holevo bound of Eq. (23), after some simple algebra we get the formula for the key rate

$$R = \log_2 \frac{2T(1 + T)}{e(1 - T)\sqrt{\sigma\sigma'}} - h(\nu_1) - h(\nu_2) + h(\bar{\nu}_1), \tag{26}$$

where ν_1 and ν_2 are given in Eqs. (14) and (15) and $\bar{\nu}_1$ is given in Eq. (21).

IV. ANALYSIS OF THE ATTACKS

Here we study the security thresholds $R = 0$ that describe the performances of the considered protocol for all possible attacks. The thresholds are given in terms of the tolerable excess noise, defined as $N := [T - 1 + \omega(1 - T)]/T$, as a function of the channel transmissivity T.

Figure 3 shows the two-way security thresholds in direct reconciliation. In particular, the (red) lines labeled (a) and (c) describe the thresholds of the two-way protocol obtained when the correlation parameters of the attack fulfill the condition $g = -g'$. In this case curve (a) describes the security threshold for maximally entangled ancillary modes E_1 and E_2. This situation is described by two distinct (although equivalent) setups of the coherent attack, for which $|g| = \sqrt{\omega^2 - 1} = -|g'|$. Curve (c), obtained when $|g| = \omega - 1 = -|g'|$, gives the extremal case of separable and maximally correlated ancillae. Black lines are the security thresholds when Eve exploits a correlation of the type $g = g'$. In this group of attacks, modes E_1 and E_2 can only share separable correlation, and for $g = \omega - 1 = g'$ we have curve (b), while for $g = 1 - \omega = g'$ we get curve (d). Finally, the dashed line provides the two-way threshold, under standard collective attacks, i.e., when $g = g' = 0$.

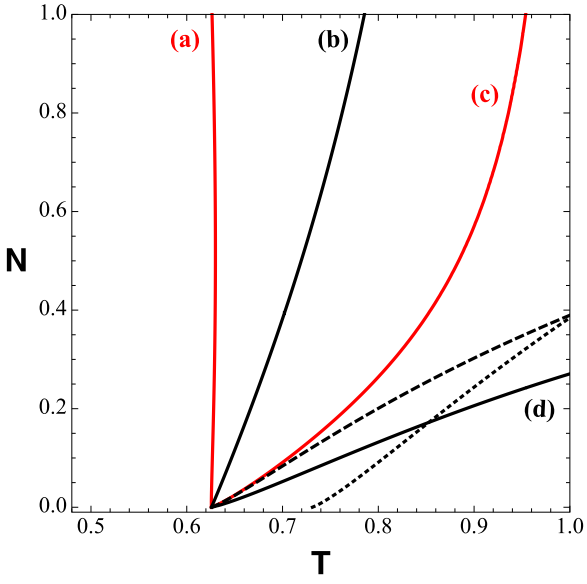

FIG. 3. (Color online) Security thresholds for the case of the two-way protocol, in direct reconciliation, against two-mode coherent attacks. Excess noise [in vacuum shot noise units (SNU)] is represented on the ordinate, and transmittivity is represented on the abscissa. Curves (a) and (c) describe two-mode attacks for which $g = -g'$. In particular, (a) is the threshold obtained when Eve uses maximally entangled ancillas E_1 and E_2. This case is given by the two equivalent conditions on the correlation parameter, $g = \sqrt{\omega^2 - 1} = -g'$ and $g = -\sqrt{\omega^2 - 1} = -g'$. Curve (c) describes the cases $g = \omega - 1 = -g'$ and $g = 1 - \omega = -g'$. Curves (b) and (d) correspond to the thresholds for $g = g'$. For curve (b) we have $g = \omega - 1 = g'$ and $g = 1 - \omega = g'$ (*d*). The dashed line is the threshold for standard collective attacks, $g = g' = 0$. The dotted black line is the security threshold for the corresponding one-way protocol, for which only collective attacks can be considered. We see that curve (d) partly goes below the one-way threshold for high transmissivities.

All these cases have been compared with the security threshold of the one-way protocol [31], in direct reconciliation (dotted line), for which the collective attacks are known to be optimal. We see that for standard collective attacks, the two-way protocol (dashed line) always overcome the performances of the one-way protocol (dotted line). However, if Eve exploits suitably correlated ancillae, she can perform a more profitable eavesdropping of the two-way protocol. This is evident from curve (d), which is clearly below the security

threshold corresponding to collective attacks (dashed line), and for high transmissivity ($T \gtrsim 0.86$), it goes below the security threshold for the one-way protocol (dotted line). Thus for the two-way protocol described in this paper, we find that the two-mode coherent attack, given by curve (d), is optimal. In the Appendix we deepen the discussion of this result.

V. CONCLUSIONS

We have studied the two-way QKD protocol, focusing on its security under two-mode coherent attacks. In the communication scheme studied here a coherent attack can be explicitly considered and analytically solved. The analysis spotlights the evidence of a coherent attack beating the collective one in the setting of point-to-point protocols.

A similar result has been obtained in previous investigations focused on the alternative approach to quantum cryptography, based on the end-to-end paradigm. As proved in Refs. [7] and [8] when the parties establish the key exploiting two channels with an untrusted middle relay, then Eve can potentially obtain an advantage by exploiting correlated ancillary modes. Here something similar happens, although the optimal attack is different [28].

Finally, our analysis confirms the importance of the on-off switching strategy, in the context of two-way QKD protocols [13]. In light of the results presented, we conclude that the active exploitation of the additional degrees of freedom available to the parties in two-way communication represents a necessary solution to avoid the possibility of powerful coherent attacks. Alice can decide to open or close the two-way quantum communication, therefore switching between one-way and two-way instances; finally, Alice and Bob decide which instances to keep on the base of Eve's strategy. In this sense the on-off switching can grant the immunity of two-way protocols against coherent attacks. Further work [24] will extend these results, here restricted to direct reconciliation, and will consider finite-size effects and composable security [29,30].

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by Leverhulme Trust and the EPSRC via qDATA (Grant No. EP/L011298/1) and the UK Quantum Communications HUB (Grant No. EP/M013472/1).

APPENDIX: OPTIMAL ATTACK

The result in Fig. 3 shows that, differently from the one-way protocol, the use of correlated ancillae is convenient for the eavesdropper. To investigate this feature further we study the behavior of the quantities defining the key rate of Eq. (26) as a function of the thermal noise ω. We fix the classical Gaussian modulation $\mu = 10^6$, for which we have verified that the asymptotic limit is largely fulfilled, and the transmissivity to the value $T = 0.65$. In the left panel in Fig. 4, we plot the mutual information I_{AB}, given in Eq. (25), and in the right panel we plot the Holevo function χ_{EA} given by Eq. (23).

First, as one would expect, we note that the mutual information (Fig. 4, left panel) decreases with increasing thermal noise. Simultaneously, Eve's accessible information, χ_{EA} (right panel), corresponding to the optimal two-mode attack,(d), is the highest among the other cases, (a)–(c). It also rapidly increases for increasing ω. This attack is profitable for Eve because she is able to increase her knowledge of Alice's variable, α, at a higher rate than Bob can. To illustrate this property further we have plotted in Fig. 5 the relative variation of Alice-Bob mutual information,

$$\Delta I_{AB} = (I_{AB} - I_c)/I_c, \tag{A1}$$

and of the Holevo function,

$$\Delta \chi_{EA} = (\chi_{EA} - \chi_c)/\chi_c, \tag{A2}$$

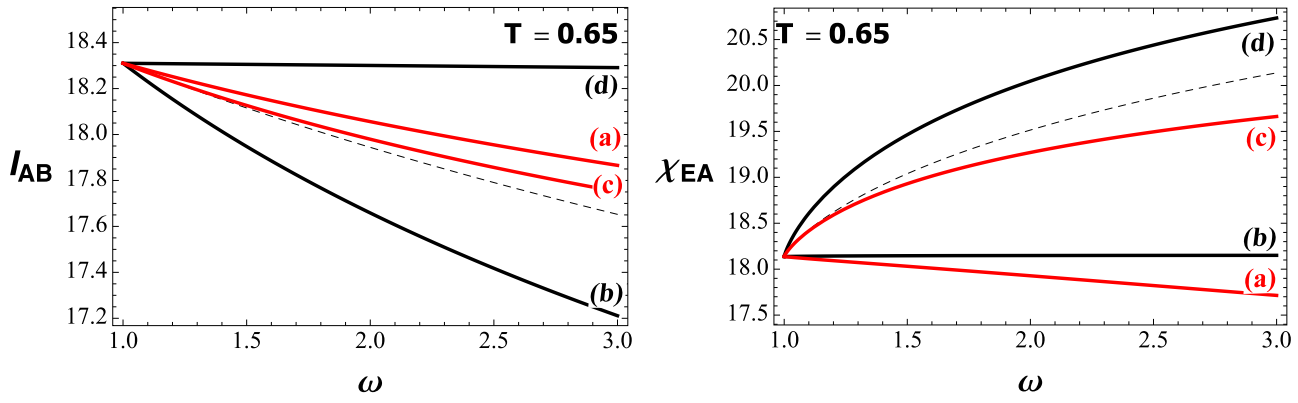

FIG. 4. (Color online) The behavior of the asymptotic mutual information I_{AB} (left) and of the Holevo function χ_{EA} (right) as a function of Eve's thermal noise ω. We fix the Gaussian modulation, $\mu = 10^6$, value for which we checked that the asymptotic limit is achieved. We also fix the transmissivity $T = 0.65$, for which the parties may obtain a positive key rate [see curves (a) and (b) in Fig. 3]. The labeling corresponds to that adopted for the thresholds in Fig. 3. We have that curve (a) describes two-mode attacks for which $g = \sqrt{\omega^2 - 1} = -g'$ or $g = -\sqrt{\omega^2 - 1} = -g'$ and curve (c) describes the case $g = \omega - 1 = -g'$ or $g = 1 - \omega = -g'$. Curve (b) corresponds to the case $g = \omega - 1 = g'$, and (d) to the case $g = 1 - \omega = g'$, i.e., the optimal attack. The dashed line refers to standard collective attacks, $g = g' = 0$. We see that, for the optimal attack (d), while the mutual information slightly decreases with increasing ω, the curve corresponding to the Holevo bound, χ_{EA}, increases and at a higher rate than in any other attack. This causes the reduction in the key rate in case (d).

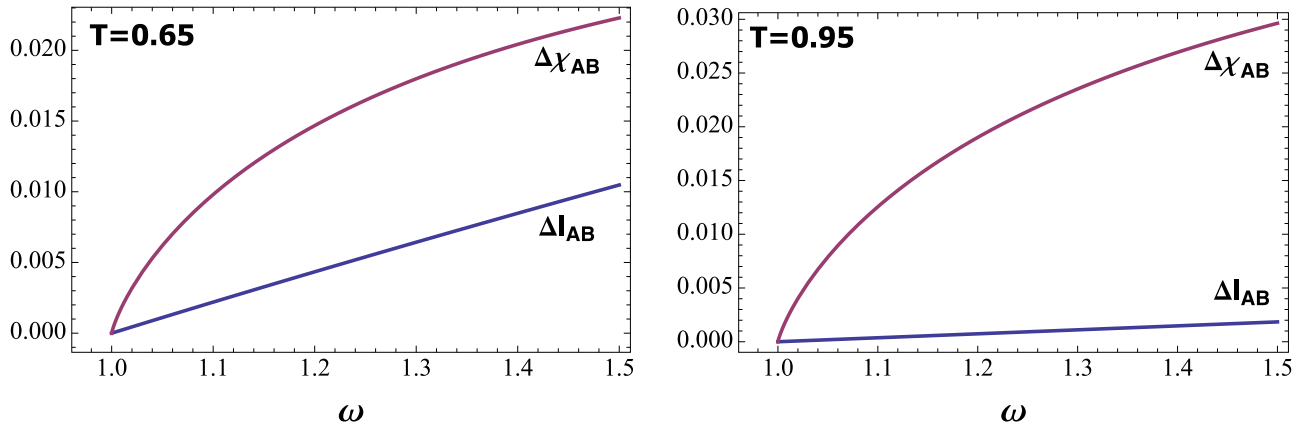

FIG. 5. (Color online) The relative variation of the Holevo bound $\Delta\chi_{EA}$ given in Eq. (A2), and of the mutual information ΔI_{EA} from Eq. (A1), for the optimal attack (d), versus ω for fixed values of the transmissivity, $T = 0.65$ (left) and $T = 0.95$ (right).

of the optimal attack with respect to the respective expressions under collective attacks ($g = g' = 0$), given by I_c and χ_c. In the left panel we plot the case for $T = 0.65$, while the right panel shows the case for $T = 0.95$. We note that with increasing transmissivity T, the relative variation in the mutual information tends to 0, while the relative variation in Eve's Holevo information tends to increase.

[1] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lutkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).

[2] W. H. Wotter and W. H. Zurek, Nature **299**, 802 (1982).

[3] C. H. Bennett and G. Brassard, *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (IEEE Press, New York 1984), pp. 175–179.

[4] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, J. Cryptol. **5**, 3 (1992).

[5] S. L. Braunstein and P. van Lock, Rev. Mod. Phys. **77**, 513 (2005).

[6] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Rev. Mod. Phys. **84**, 621 (2012).

[7] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Gehring, C. S. Jacobsen, and U. L. Andersen, Nature Phot. **9**, 397 (2015). See also arXiv:1312.4104.

[8] C. Ottaviani, G. Spedalieri, S. L. Braunstein, and S. Pirandola, Phys. Rev. A **91**, 022320 (2015).

[9] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Gehring, C. S. Jacobsen, and U. L. Andersen, Nature Phot. **9**, 773 (2015); arXiv:1506.06748.

[10] G. Spedalieri, C. Ottaviani, S. L. Braunstein, T. Gehring, C. S. Jacobsen, U. L. Andersen, and S. Pirandola, in *Proceedings of the SPIE Security + Defence 2015 Conference on Quantum Information Science and Technology, Toulouse, France (21–24 September 2015)* (2015), paper 9648-47.

[11] F. Grosshans, G. Van Assche, J. Wenger, R. Tualle-Brouri, N. J. Cerf, and P. Grangier, Nature **421**, 238 (2003).

[12] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, Phys. Rev. Lett. **93**, 170504 (2004).

[13] S. Pirandola, S. Mancini, S. Lloyd, and S. L. Braunstein, Nature Phys. **4**, 726 (2008).

[14] C. Weedbrook, C. Ottaviani, and S. Pirandola, Phys. Rev. A **89**, 012309 (2014).

[15] M. Lucamarini and S. Mancini, Phys. Rev. Lett. **94**, 140501 (2005).

[16] J. H. Shapiro, Phys. Rev. A **80**, 022320 (2009).

[17] Z. Zhang, M. Tengner, T. Zhong, F. N. C. Wong, and J. H. Shapiro, Phys. Rev. Lett. **111**, 010501 (2013).

[18] Q. Zhuang, Z. Zhang, J. Dove, F. N. C. Wong, and J. H. Shapiro, arXiv:1508.01471.

[19] S. Lloyd, Science **321**, 1463 (2008).

[20] S.-H. Tan *et al.*, Phys. Rev. Lett. **101**, 253601 (2008).

[21] S. Barzanjeh *et al.*, Phys. Rev. Lett. **114**, 080503 (2015).

[22] S. Pirandola, New J. Phys. **15**, 113046 (2013).

[23] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, Rev. Mod. Phys. **86**, 1203 (2014).

[24] C. Ottaviani and S. Pirandola (unpublished).

[25] R. Renner and J. I. Cirac, Phys. Rev. Lett. **102**, 110504 (2009).

[26] S. Pirandola, S. L. Braunstein, and S. Lloyd, Phys. Rev. Lett. **101**, 200504 (2008).

[27] S. Pirandola, C. Ottaviani, C. S. Jacobsen, G. Spedalieri, S. L. Braunstein, T. Gehring, and U. L. Andersen (unpublished); arXiv:1505.07457.

[28] In end-to-end QKD based on untrusted relays (performing a CV Bell detection) the optimal attack, at fixed thermal noise ω, is one of the two of type (a).

[29] A. Leverrier, Phys. Rev. Lett. **114**, 070501 (2015).

[30] E. Diamanti and A. Leverrier, Entropy **17**, 6072 (2015).

[31] In Fig. 3 we want to show not only that the optimal two-mode attack can reduce the security threshold of the two-way protocol below that against collective attacks, but that even for large enough T, the optimal two-mode threshold is lower than the one-way protocol against collective.