
Article:
Papanastasiou, Panagiotis, Ottaviani, Carlo orcid.org/0000-0002-0032-3999 and Pirandola, Stefano orcid.org/0000-0001-6165-5615 (2017) Finite-size analysis of measurement-device-independent quantum cryptography with continuous variables. Physical Review A. 042332. ISSN 1094-1622

https://doi.org/10.1103/PhysRevA.96.042332

Finite-size analysis of measurement-device-independent quantum cryptography with continuous variables

Panagiotis Papanastasiou, Carlo Ottaviani, and Stefano Pirandola
Computer Science and York Centre for Quantum Technologies, University of York, York YO10 5GH, United Kingdom
(Received 14 July 2017; published 23 October 2017)

We study the impact of finite-size effects on the key rate of continuous-variable (CV) measurement-device-independent (MDI) quantum key distribution (QKD), considering two-mode Gaussian attacks. Inspired by the parameter estimation technique developed in by Ruppert *et al.* [Phys. Rev. A **90**, 062310 (2014)], we adapt it to study CV-MDI-QKD and, assuming realistic experimental conditions, we analyze the impact of finite-size effects on the key rate. We find that the performance of the protocol approaches the ideal one, increasing the block size, and, most importantly, that blocks between 10^6 and 10^9 data points may provide key rates $\sim 10^{-2}$ bit/use over metropolitan distances.

DOI: 10.1103/PhysRevA.96.042332

I. INTRODUCTION

Quantum key distribution (QKD) [1] promises to allow unconditionally secure (theoretical) communication. Its strength relies on two main elements: The encoding of classical information (0 and 1 bits) into nonorthogonal quantum states, and the impossibility of perfect discrimination between them. In a conventional QKD protocol two users, Alice and Bob, share quantum systems which are unavoidably corrupted every time that an eavesdropper (Eve) tries to access the information encoded. This perturbation is detectable and allows the parties to quantify the amount of error correction and privacy amplification to apply to the shared data, in order to reduce Eve's information to a negligible amount. Then they can use the obtained key in an one-time pad protocol [2].

The fundamental mechanism of QKD is clearly preserved also in more complex (repeater-based) communication configurations [3,4], aiming at activating long-distance communication and quantum networks [5,6]. In the basic point-to-point scenario, the recent work [7] succeeded in establishing the secret-key capacity of various quantum channels. The combined use of relative entropy of entanglement [8–10] and teleportation stretching (which reduces any adaptive protocol to a block form) enables one to compute the two-way capacity of many important quantum channels (see Ref. [7] and further works [11–15] for the correct definition and rigorous use of teleportation stretching in quantum communication, quantum metrology, and channel discrimination). The result of Ref. [7] sets the fundamental limit of point-to-point QKD and, as such, it marks the edge when private communication inevitably needs quantum repeaters. This benchmark has already had a wide application in recent works [16–29].

Continuous-variable (CV) quantum systems [30], in particular Gaussian systems [31], emerged recently as very promising carriers of quantum information. They have the potential to be used for high-rate quantum communication because, rather than using single-photon quantum states and photon counting, they employ bright coherent states and homodyne detections, which naturally boost the achievable key rate. Based on this premise, CV-QKD protocols [32] have been proposed using one-way [33–35] or two-way quantum communications [36]. Some one-way schemes have been experimentally realized [37–40], over remarkably long distances [41,42]. Additional theoretical analysis has been focused on QKD with thermal states [43–48], with an experiment performed [49]. Recently, CV-QKD has been extended to a network configuration [50,51], implementing the general idea of measurement-device-independent (MDI) QKD [52,53]. Here two parties, unable to access a secure direct link, can be assisted by an intermediate relay (even untrusted) to establish a secure channel.

Many challenges [54] need to be solved before private quantum networking can become a mature technology. However, CV-QKD protocols and their security analysis have progressed rapidly toward more practical and realistic assumptions. In this respect, the incorporation of finite-size effects is particularly important. In fact, when we assume that the parties exchange only a finite number of signals, one expects the deterioration of the key rate. In addition to this, finite-size analysis is also the first step toward a more general security proof within the composable framework [55–57]. While the theoretical study of the impact of finite-size effects has been done in several previous works [58–60], CV-MDI QKD has been so far investigated neglecting this aspect and limiting the analysis to the asymptotic regime [61].

In order to start filling this gap, in this work we focus on evaluating the impact of finite-size effects on the key rate of a CV-MDI protocol. This study is important not only because these effects have not yet evaluated in detail for MDI protocols, but also because this type of analysis represents a necessary step to refine the security analysis of CV-MDI toward the more complete composable scenario. We perform a detailed study of the impact of finite-size effects for both the symmetric [51] and asymmetric [50,62] configurations. We extend the parameter estimation methods described in Ref. [60], for conventional one-way protocols, to relay-based communication. We consider Gaussian two-mode attacks which have been already extensively studied in one-way schemes [63], two-way protocols [64,65], and the CV-MDI setup [66].

We remark that, in our analysis, we work within the Gaussian assumption. This allows us to develop the statistical estimation theory of the relevant parameters of the channels, which are their transmissivities and excess noise. The confidence interval of the estimated parameters are then quantified using their variances and setting a 6.5-σ accuracy, which

allows us to grant a very low error probability of $\epsilon_{PE} = 10^{-10}$ during the parameter estimation procedure. The confidence intervals are used to select the worst-case scenario, choosing lower transmissivity and higher excess noise in the links.

In order to compute the key rate, we explicitly include a correction term Δ which accounts for the use of the Holevo function (which is an asymptotic quantity) in the finite-size regime [60]. The key rate is then numerically computed, using the estimated values of transmissivity and noise, and optimized over free parameters, which are the Gaussian modulation of the signals and the ratio between the number of signals used in the parameter estimation and total number of signals exchanged.

As expected, we find that, by increasing the block-size of the signals exchanged, one recovers the performance under ideal condition. Most importantly, one has that block size in the range of $10^6 \div 10^9$ signals can provide a positive key rate of about 10^{-2} bit/use, in the presence of high excess noise of 0.01 vacuum shot noise units (SNU) and attenuation compatible with the use of standard optical fibers over metropolitan distances. The structure of this paper is the following: In Sec. II, we present the details of the CV-MDI-QKD protocol. In Sec. III, we describe the parameter estimation. In Sec. IV, we discuss the results obtained, and finally Sec. V is for the conclusions.

II. PROTOCOL, EAVESDROPPING, AND KEY RATE

For the sake of clarity, let us first describe the working mechanism of CV-MDI QKD from the prepare and measure perspective, where Alice and Bob send coherent states, $|\alpha\rangle$ and $|\beta\rangle$, to an intermediate relay. The amplitudes α and β are Gaussian modulated; i.e., each party sends to the relay an average thermal state with variance $V_M \geqslant 0$. The duty of the relay is very simple: It mixes the incoming signals on a balanced beam splitter and performs a CV Bell detection, i.e., two conjugate homodyne detections on q_- and p_+, at the output ports of the balanced beam splitter [67].

Then, the relay broadcasts the obtained values of $\gamma := (q_- + ip_+)/\sqrt{2}$. This new variable can also be written as $\gamma := \alpha - \beta^* + \hat{\delta}$, where $\hat{\delta}$ is the detection noise. It is then clear that the relay acts as a correlator for the parties, who can each infer the other variable (α, β) from a simple postprocessing [50].

The broadcast of the Bell detection outcomes, γ, does not help the eavesdropper who is forced to attack the communication links to the relay in order to obtain information on amplitudes α and β. This operation introduces detectable excess of noise that the parties can use to quantify Eve's knowledge on α and β (accessible information). From this stage on, the protocol works as any other QKD scheme [1], with the Alice and Bob implementing enough error correction and privacy amplification to reduce Eve's accessible information to a negligible amount.

A. Two-mode eavesdropping

A powerful approach to study the security of any quantum cryptographic protocol is to adopt the entanglement-based (EB) representation, where the description of the dynamics takes place in a dilated Hilbert space, which allows us to work with pure states. The EB representation of CV-MDI QKD scheme is given in Fig. 1: Alice's and Bob's sources of coherent states are purified assuming to start from a two-mode squeezed vacuum (TMSV) states ρ_{aA} and ρ_{bB}, whose zero-mean Gaussian states are completely described by the following identical covariance matrices (CM)

$$\mathbf{V}_{aA} = \mathbf{V}_{bB} = \begin{pmatrix} \mu\mathbf{I} & \sqrt{\mu^2-1}\mathbf{Z} \\ \sqrt{\mu^2-1}\mathbf{Z} & \mu\mathbf{I} \end{pmatrix}, \tag{1}$$

where $\mu = V_M + 1$ and $\mathbf{Z} = \mathrm{diag}(1,-1)$.

Modes A and B are sent through the links, while local ones, a and b, are heterodyned. The measurements projects the traveling modes into coherent states $|\alpha\rangle$ and $|\beta\rangle$ respectively. The channel attenuation on modes A and B is modeled by two beam splitters with transmissivities τ_A and τ_B, with $0 \leqslant \tau_{A,B} \leqslant 1$. These process Alice's and Bob's signals with a pair of Eve's ancillary systems E_1 and E_2 which, in general, belong to a wider reservoir of modes controlled by the eavesdropper and including the set $\mathbf{e}$ (which can be neglected in the limit of infinite signals exchanged [68]).

We can then write the dilation of the initial Eve's state as a two-mode Gaussian state $\sigma_{E_1E_2}$ described by the following general CM:

$$\mathbf{V}_{E_1E_2} = \begin{pmatrix} \omega_A\mathbf{I} & \mathbf{G} \\ \mathbf{G} & \omega_B\mathbf{I} \end{pmatrix}, \tag{2}$$

where $\mathbf{G} = \mathrm{diag}(g, g')$. The correlation parameters g and g' satisfy the constraints given in Ref. [69], while $\omega_A, \omega_B \geqslant 1$ account for the thermal noise injected by Eve, on each link, during the attack. When $g = g' = 0$, the two-mode state $\sigma_{E_1E_2}$ is a tensor product, which leads to a standard single-mode collective attack realized by two independent entangling cloners [38]. By contrast for $g \neq 0$ and $g' \neq 0$, the

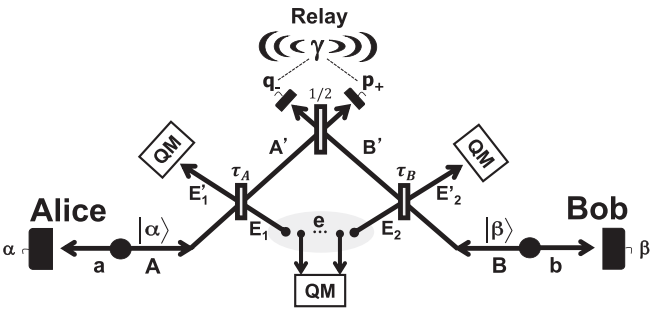

FIG. 1. The figure shows the EB representation of CV-MDI QKD. Alice and Bob have TMSV states with modes (a, A) and (b, B). Local modes a and b are kept by the parties, while A and B are sent to the relay through two links with transmittance τ_A and τ_B. When Alice and Bob heterodyne the local modes, the traveling ones A and B are projected onto coherent states $|\alpha\rangle$ and $|\beta\rangle$. The relay performs a Bell measurement and broadcasts the outcomes γ, creating correlation between the parties: For instance, Bob recovers Alice variable β by subtracting his variable α from the relay's outputs γ. The Gaussian attack on the links is simulated by Eve using ancillas E_1 and E_2 and thermal noise $\omega_A \geqslant 1$ and $\omega_B \geqslant 1$, respectively. These ancillary modes are, in general, two-mode correlated (see text for more details). The ancillary outputs are stored in a quantum memory for a later measurement.

two entangling cloners are not independent, and the optimal attack is two-mode coherent, as described in Refs. [50,51,66].

B. Key rate

The EB representation is useful for the security analysis because Alice-Bob reduced output states $\rho_{ab|\gamma}$ and $\rho_{b|\gamma\alpha}$, described by the CMs $\mathbf{V}_{ab|\gamma}$ and $\mathbf{V}_{b|\gamma\alpha}$ respectively (see Appendix A for further details), have the same entropies of Eve's output states. Under the ideal assumption that the parties exchange infinitely many signals ($N \gg 1$), and assuming that the parties reconcile over Alice's data to build the key, one bounds Eve's accessible information by the Holevo function,

$$I_H := S(\rho_{ab|\gamma}) - S(\rho_{b|\gamma\alpha}), \tag{3}$$

where $S(.)$ is the von Neumann entropy. For Gaussian states, we have the simple expression [31]

$$S = \sum_x h(x),$$

where x is the generic symplectic eigenvalue of the CM, and

$$h(x) = \frac{x+1}{2}\log_2\frac{x+1}{2} - \frac{x-1}{2}\log_2\frac{x-1}{2}, \tag{4}$$

$$\overset{x\to\infty}{\to} \log_2\frac{e}{2}x. \tag{5}$$

We then can write an expression for the key rate

$$K^\infty := \xi I_{AB} - I_H, \tag{6}$$

where $\xi \leqslant 1$ quantifies the inefficiency of error correction and privacy amplification protocols [70–72] and I_{AB} is Alice-Bob mutual information. This is given by

$$I_{AB} = \frac{1}{2}\log_2\frac{V^q_{b|\gamma}+1}{V^q_{b|\gamma\alpha}+1} + \frac{1}{2}\log_2\frac{V^p_{b|\gamma}+1}{V^p_{b|\gamma\alpha}+1}, \tag{7}$$

with $V^q_{b|\gamma}$ ($V^p_{b|\gamma}$) and $V^q_{b|\gamma\alpha}$ ($V^p_{b|\gamma\alpha}$) being the variances of CMs $\mathbf{V}_{ab|\gamma}$ and $\mathbf{V}_{b|\gamma\alpha}$ for the position (momentum) quadrature. These CMs are given in Appendix A. The key rate is then function of parameters $\xi, \omega_A, \omega_B, \tau_A$, and τ_B and the Gaussian modulation V_M. Its expression can be found in the supplemental information of Ref. [50].

III. CHANNEL PARAMETER ESTIMATION

In a practical implementation of any QKD protocol, Alice and Bob can only exchange a finite number of signals. In addition, they can only use a portion of these to build the key, because the others are used to estimate the channel parameters. In this section, we provide a description of CV-MDI QKD, quantifying the impact of finite-size effects and the performance of the protocol. To perform this analysis, we adapt the theory developed in Ref. [60] for one-way CV QKD. We determine the channel parameters (transmissivity and excess noise) within confidence intervals. Then we choose the worst-case scenario, picking the lower transmissivity and higher excess noise within their confidence intervals, so as to minimize the key rate.

A. Losses and excess noise at the relay outputs

The outputs variables of the relay are quadratures q_-, relative to mode $-$, and p_+ for mode $+$. These depend on the evolution of Alice's and Bob's traveling modes $A = (q_A, p_A)$ and $B = (q_B, p_B)$. In terms of these input field quadratures, one can then write the following relations:

$$q_- = \frac{1}{\sqrt{2}}(\sqrt{\tau_B}q_B - \sqrt{\tau_A}q_A) + q_N, \tag{8}$$

$$p_+ = \frac{1}{\sqrt{2}}(\sqrt{\tau_B}p_B + \sqrt{\tau_A}p_A) + p_N, \tag{9}$$

where $q_N = q_\epsilon + q_{sn}$ and $p_N = p_\epsilon + p_{sn}$ are noise terms accounting for both excess noise and quantum shot noise coming from the signal modes as well as Eve's ancillary modes. Their variances are given by

$$V_{q_N} = 1 + V_{q,\epsilon}, \quad V_{p_N} = 1 + V_{p,\epsilon}, \tag{10}$$

with

$$V_{q,\epsilon} = k - gu, \quad V_{p,\epsilon} = k + g'u, \tag{11}$$

and

$$k = \frac{(1-\tau_B)(\omega_B - 1) + (1-\tau_A)(\omega_A - 1)}{2}, \tag{12}$$

$$u = \sqrt{(1-\tau_B)(1-\tau_A)}, \tag{13}$$

where g and g' have been defined in Eq. (2).

Now we describe in more detail the parameter estimation procedure. Alice and Bob's Gaussian modulation V_M is assumed to be a known parameter. We need to estimate the channel's transmissivity τ_A, τ_B and variance of the excess noises $V_{q,\epsilon}$ and $V_{p,\epsilon}$, with their confidence intervals. Assuming that m Gaussian distributed signals are used for this task, we associate to $A_{q,i}$ ($A_{p,i}$) and $B_{q,i}$ ($B_{p,i}$), for $i \in \{1,2,\ldots,m\}$, the empirical realizations of the field quadrature of the traveling modes. By contrast, we denote by $R_{q,i}$ and $R_{p,i}$ the realizations of the relay outputs. Let first discuss the dynamics of the quadrature q. From Eq. (8) one can write the estimator of transmissivity τ_A as follows:

$$\hat{\tau}_{Aq} = \frac{2}{V_M^2}\hat{C}^2_{AR_q},$$

where the covariance $C_{AR_q} = \sqrt{\tau_A/2}V_M$ has maximum likelihood estimator given by

$$\hat{C}_{AR_q} = \frac{1}{m}\sum_{i=1}^m A_{q,i}R_{q,i},$$

and to which one can associate the following variance (see Appendix C for more details):

$$\mathrm{Var}(\hat{\tau}_{Aq}) = \frac{8\tau_A}{m}\left(\tau_A + \frac{\tau_B}{2}\right)\left[1 + \frac{V_{q,N}}{\left(\tau_A + \frac{\tau_B}{2}\right)V_M}\right]. \tag{14}$$

Very similar relations hold for the estimator of τ_A obtained considering the other output of the relay, p_+. We can write the estimator of the covariance C_{AR_p}, which is given by

$$\hat{C}_{ARp} = \frac{1}{m}\sum_{i=1}^m A_{p,i}R_{p,i},$$

and then using Eq. (9) one can write the estimator of the transmissivity τ_A

$$\hat{\tau}_{Ap} = \frac{2}{V_M^2}\hat{C}^2_{AR_p},$$

having variance

$$\mathrm{Var}(\hat{\tau}_{Ap}) = \frac{8}{m}\tau_A\left(\tau_A + \frac{\tau_B}{2}\right)\left[1 + \frac{V_{p,N}}{\left(\tau_A + \frac{\tau_B}{2}\right)V_M}\right]. \quad (15)$$

We notice that it differs from the formula of Eq. (14) for the expression of $V_{p,N}$, given in Eq. (10). Now, from Eqs. (14) and (15), we calculate the optimum linear combination of the variances of the two estimators:

$$\mathrm{Var}(\hat{\tau}_A) = \frac{\mathrm{Var}(\hat{\tau}_{A_q})\mathrm{Var}(\hat{\tau}_{Ap})}{\mathrm{Var}(\hat{\tau}_{A_q}) + \mathrm{Var}(\hat{\tau}_{Ap})} := \sigma_A^2. \quad (16)$$

The same steps can be performed to obtain the relevant estimators for transmissivity τ_B and the corresponding variance $\mathrm{Var}(\hat{\tau}_B) = \sigma_B^2$.

Now we write the estimator of the variance of the excess noise present on the communication links, $V_{q,\epsilon}$. This can be derived from the maximum likelihood estimator for $V_{q,N}$, and it reads

$$\hat{V}_{q,\epsilon} = \frac{1}{m}\sum_{i=1}^{m}\left[R_{q,i} - \frac{1}{\sqrt{2}}(\sqrt{\hat{\tau}_B}B_{q,i} - \sqrt{\hat{\tau}_A}A_{q,i})\right]^2 - 1, \quad (17)$$

with variance (Appendix C)

$$\mathrm{Var}(\hat{V}_{q,\epsilon}) \approx \frac{2}{m}V_{q,N}^2 := s_q^2. \quad (18)$$

Correspondingly, we obtain an estimator for $V_{p,\epsilon}$ expressed as

$$\hat{V}_{p,\epsilon} = \frac{1}{m}\sum_{i=1}\left[R_{p,i} - \frac{1}{\sqrt{2}}(\sqrt{\hat{\tau}_B}B_{p,i} + \sqrt{\hat{\tau}_A}A_{p,i})\right]^2 - 1 \quad (19)$$

and variance

$$\mathrm{Var}(\hat{V}_{p,\epsilon}) \approx \frac{2}{m}V_{p,N}^2 := s_p^2. \quad (20)$$

Finally, from Eqs. (17) and (20) we compute the confidence intervals and select the pessimistic values given by the following choice of parameters:

$$\tau_A^{\mathrm{low}} = \tau_A - 6.5\sigma_A, \quad \tau_B^{\mathrm{low}} = \tau_B - 6.5\sigma_B, \quad (21)$$

$$V_{q,\epsilon}^{\mathrm{up}} = V_{q,\epsilon} + 6.5s_q, \quad V_{p,\epsilon}^{\mathrm{up}} = V_{p,\epsilon} + 6.5s_p. \quad (22)$$

B. Secret key rate with finite size effects

Once we have obtained the estimation of the transmissivities of the links and the corresponding excess noises, we can write the key rate, incorporating finite-size effects writing

$$K = \frac{n}{\bar{N}}\left[K^{\infty}\left(\xi, V_M, \tau_A^{\mathrm{low}}, \tau_B^{\mathrm{low}}, V_{q,\epsilon}^{\mathrm{up}}, V_{p,\epsilon}^{\mathrm{up}}\right) - \Delta(n)\right], \quad (23)$$

where $n = \bar{N} - m$ is the number of signals used to prepare the key and $\bar{N}$ is the total number of signals exchanged. The key rate is then computed replacing the values of Eqs. (21) and (22) in the asymptotic key rate of Eq. (6). In particular, the first one computes the key rate R of Eq. (23), using the Holevo function of Eq. (3) for the channel parameters given by Eqs. (21) and (22). Then, in order to account for the penalty for using the Holevo function even if we have a finite number of signals exchanged, one must include the correction term

$$\Delta(n) \sim \sqrt{\frac{1}{n}\log_2 2\varepsilon_{PA}^{-1}},$$

which depends on the number of signals used to prepare the key, n, and the probability of error related to the privacy amplification procedure ε_{PA}. A detailed description of this correction term can be found in Ref. [59].

IV. RESULTS

The key rate of the asymmetric configuration of the relay is described in Fig. 2(a). We plot the key rate as a function of Bob's channel transmissivity, expressed in terms of dB of attenuation, while the transmissivity of Alice's link is set to $\tau_A = 0.98$. The curves are obtained considering two-mode optimal attacks, for which $g = -g'$ with $g = \min[\sqrt{(\omega_A - 1)(\omega_B + 1)}, \sqrt{(\omega_B - 1)(\omega_A + 1)})]$ and $\omega_A \sim \omega_B \sim 1.01$ [50] and using the key rate of Eq. (23) incorporating also finite-size effects. The efficiency of classical code for error correction and reconciliation efficiency is set to $\xi = 0.98$, and the final key rate is optimized over the variance of the Gaussian modulation V_M (see Fig. 3) and the ratio $r = n/\bar{N}$. The black solid line gives the asymptotic key rate for very large $\bar{N}$ ($>> 10^9$), while the dashed line is for block size $\bar{N} = 10^9$ and the dot-dashed line is obtained for $\bar{N} = 10^6$.

The bottom panel in Fig. 2(b) plots the secret key rate for the symmetric case [51] ($\tau_A = \tau_B$). The curves are obtained setting all the other parameters as in Fig. 2(a) and optimizing the key rate as before for the case including finite-size effects. The black solid line describes again the asymptotic case $\bar{N} \to \infty$ of the symmetric configuration while the dashed lines is obtained for $\bar{N} = 10^9$ and the dotted one for $\bar{N} = 10^6$.

Let us finally remark on a couple of points. First, we notice that the performance of finite-size CV-MDI-QKD converges to the ideal one if the number of signals exchanged increases. Second, according to the plots, we notice that the key rate of the CV-MDI-QKD protocol is sufficiently robust with respect to the finite-size effects, with block sizes of 10^9 points approaching the asymptotic limit.

V. CONCLUSION

In conclusion, we have studied the security of Gaussian CV-MDI QKD, taking into account finite-size effects. These emerge when one assumes that the parties exchange only a finite number of signals during the quantum communication stage. In our analysis, we assumed imperfect efficiency of error correction and privacy amplification ($\xi < 1$) and developed the finite-size analysis adapting the channel parameters estimation approach described in Ref. [60]. The resulting finite-size key rate has then been optimized over the Gaussian modulation and the number of signals used to perform the parameter estimation.

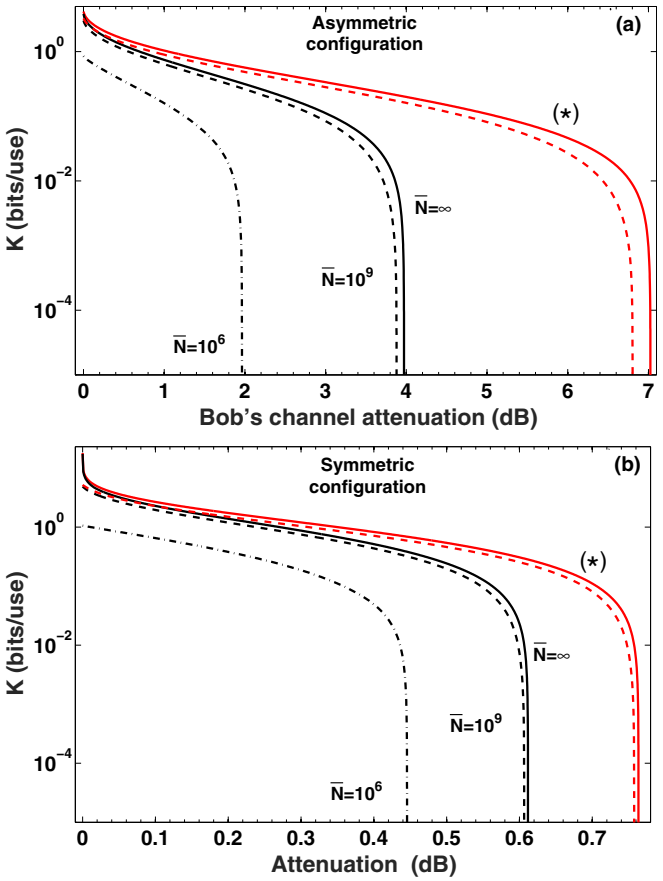

FIG. 2. The figure summarizes the impact of finite-size effects on the performance of CV-MDI QKD for both asymmetric (a) and symmetric (b) configurations of the relay, in the presence of optimal two-mode attack. In panel (a), the key rate is plotted as a function of the dB of attenuation on Bob's channel, with the relay placed near Alice $\tau_A = 0.98$. From top to bottom, the black curves describe the rate for $\bar{N} \gg 1$ with $\xi = 0.98$ and optimizing over V_M (solid line). Then we have the cases with finite block size. The dashed line is for $\bar{N} = 10^9$ while the dot-dashed curve is obtained for $\bar{N} = 10^6$. In all cases, the excess noise is about 0.01 SNU. The red curves (*) describe the case obtained for pure loss and assuming $\xi = 1$, $V_M \to \infty$, $\bar{N} \to \infty$ (solid line) and $\bar{N} = 10^9$ (dashed line). Panel (b) focuses on the symmetric configuration of the relay. The curves are obtained using the same parameters as in panel (a), but setting $\tau_A = \tau_B = \tau$.

Our results show that when also considering finite-size effects under realistic conditions, CV-MDI QKD over metropolitan distances is feasible within today's state-of-the-art experimental conditions. In particular, we found that the adoption of block size in the range $\bar{N} = 10^6 \div 10^9$ is already sufficient in order to achieve a high key rate of 10^{-2} bits/use over metropolitan distances, and in the presence of an excess noise of about 0.01 SNU.

Finally, we underline that the present analysis is not the final word on the performances of finite-size CV-MDI QKD. The validity of the described analysis is in fact restricted to the case of Gaussian attacks. Further studies are needed where finite-size effects are investigated within the composable security framework.

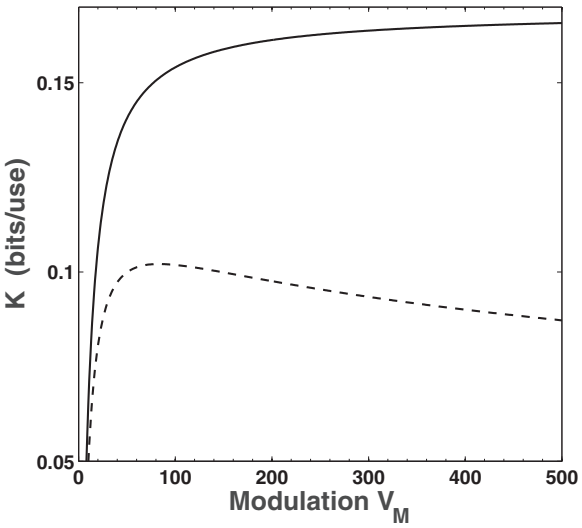

FIG. 3. This figure shows the impact of the reconciliation efficiency on the key rate. When $\xi = 0.95$ (dashed line), the Gaussian modulation maximizing the key rate is $V_M < \infty$. While when $\xi = 1$ then the optimal key rate is obtained for $V_M \to \infty$ (solid line). The lines are obtained for pure loss attack, $\tau_A = 0.98$ and $\tau_B = 0.7$, and block size $\bar{N} = 10^6$.

See Ref. [73] for a fully composable security proof of CV-MDI-QKD.

Note added. Recently, an independent work [74] has been posted on the arXiv. This work also studies the impact of finite-size blocks on the key rate of CV-MDI QKD.

ACKNOWLEDGMENTS

C.O. acknowledges Cosmo Lupo for useful discussions. This work has been supported by the EPSRC via the UK Quantum Communications HUB (Grant No. EP/M013472/1).

APPENDIX A: COVARIANCE MATRICES AND SYMPLECTIC EIGENVALUES

For the sake of clarity, in this section we rewrite the CMs and the relevant symplectic eigenvalues derived in Ref. [50] using the notation adopted in the main text. The CM describing the total output state of Alice and Bob, after the relay measurements, $\rho_{ab|\gamma}$, is given by the following expression:

$$\mathbf{V}_{ab|\gamma} = \begin{pmatrix} (V_M+1)\mathbf{I} & 0 \\ 0 & (V_M+1)\mathbf{I} \end{pmatrix} - V_M(V_M+2) \times \begin{pmatrix} \frac{\tau_A}{\varphi} & 0 & -\frac{\sqrt{\tau_A\tau_B}}{\varphi} & 0 \\ 0 & \frac{\tau_A}{\varphi'} & 0 & \frac{\sqrt{\tau_A\tau_B}}{\varphi'} \\ -\frac{\sqrt{\tau_A\tau_B}}{\varphi} & 0 & \frac{\tau_B}{\varphi} & 0 \\ 0 & \frac{\sqrt{\tau_A\tau_B}}{\varphi'} & 0 & \frac{\tau_B}{\varphi'} \end{pmatrix}, \quad \text{(A1)}$$

where

$$\varphi = (\tau_A + \tau_B)V_M + 2 + 2V_{\epsilon,g}, \quad \text{(A2)}$$

$$\varphi' = (\tau_A + \tau_B)V_M + 2 + 2V_{\epsilon,g'}, \quad \text{(A3)}$$

and where we have defined

$$V_{\epsilon,g} = \tfrac{1}{2}\big(\bar{\tau}_B(\omega_B - 1) + \bar{\tau}_A(\omega_A - 1) - 2g\sqrt{\bar{\tau}_B\bar{\tau}_A}\big), \quad \text{(A4)}$$

$$V_{\epsilon,g'} = \tfrac{1}{2}\big(\bar{\tau}_B(\omega_B - 1) + \bar{\tau}_A(\omega_A - 1) + 2g'\sqrt{\bar{\tau}_B\bar{\tau}_A}\big), \quad \text{(A5)}$$

with $\bar{\tau}_l = 1 - \tau_l$, for $l = A, B$.

Bob's output CM after the double conditioning, first with respect the relay measurements and then after Alice's heterodyne detection, $\mathbf{V}_{ab|\gamma}$, is given by

$$\mathbf{V}_{b|\gamma\alpha} = \begin{pmatrix} \frac{2(V_M+1)(1+V_{\epsilon,g})-\tau_B V_M}{2(1+V_{\epsilon,g})+\tau_B V_M} & 0 \\ 0 & \frac{2(V_M+1)(1+V_{\epsilon,g'})-\tau_B V_M}{2(1+V_{\epsilon,g'})+\tau_B V_M} \end{pmatrix},$$

which has the following symplectic eigenvalue given by the $\sqrt{\cdot}$ of the determinant of previous matrix

$$\bar{\nu} = \sqrt{\det \mathbf{V}_{b|\gamma\alpha}}. \quad \text{(A6)}$$

APPENDIX B: USEFUL ELEMENTS OF ESTIMATION THEORY

According to the method of maximum likelihood estimation, for a bivariate normal distribution $X = (X_1, X_2)$, the estimators for the mean $\mu = (\mu_1, \mu_2)$ and the covariance matrix $\mathbf{V}$ are given by

$$\hat{\boldsymbol{\mu}} = \frac{1}{m}\sum_{i=1}^{m} \mathbf{X}_i, \quad \text{(B1)}$$

$$\widehat{\mathbf{V}} = \frac{1}{m}\sum_{i=1}^{m} (\mathbf{X}_i - \hat{\boldsymbol{\mu}})(\mathbf{X}_i - \hat{\boldsymbol{\mu}})^T, \quad \text{(B2)}$$

where $\mathbf{X}_i$ is the ith statistical realization out of m realizations of $\mathbf{X}$.

The central limit theorem states that, assuming m realizations $X_1, X_2, \ldots, X_m$ ($m \gg 1$) of a random variable X with unknown density function f, mean μ, and variance $\sigma^2 < \infty$, the sample mean

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i \quad \text{(B3)}$$

is approximately normal with mean μ and variance σ^2/m. In order to estimate the mean value of a variable Y, that depends on the square of a variable X for which we have m realizations, we can use the following result: For m realizations X_i, for $i = 1, 2, \ldots, m$, of a normally distributed variable X, having mean μ and unit variance, the variable

$$Y = \sum_{i=1}^{m} X_i^2 \sim \chi^2(k, \lambda) \quad \text{(B4)}$$

is distributed according to the χ^2 distribution with $k = m$ degrees of freedom and $\lambda = m\mu^2$. The mean value and variance of the χ^2 distribution is given by

$$\mathbb{E}(Y) = k + \lambda \quad \text{(B5)}$$

and

$$\mathrm{Var}(Y) = 2(k + 2\lambda). \quad \text{(B6)}$$

Let us assume to have two estimators $\hat{s}_1$ and $\hat{s}_2$, with variances σ_1^2 and σ_2^2, for the same quantity s acquired by different processes. We then compute the optimal linear combination of the variances by the following formula:

$$\sigma^2_{\mathrm{opt}} = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}. \quad \text{(B7)}$$

APPENDIX C: VARIANCES OF THE CHANNEL PARAMETER ESTIMATORS

Let us suppose that $A_{q,i}$ ($B_{q,i}$) are independent variables, each one following the normal distribution q_A (q_B) with zero mean and variance V_M as described in Subsec. III A. Accordingly, $R_{q,i}$ ($R_{p,i}$) are assumed to be independent variables following the normal distribution of q_R (p_R), i.e., the relay output variable.

1. Variance of the transmissivity

For the covariance between modes A and R, we can write the following estimator,

$$\hat{C}_{AR_q} = \frac{1}{m}\sum_{i=1}^{m} A_{q,i} R_{q,i}, \quad \text{(C1)}$$

normally distributed as the sample mean of variable $Z = A_q R_q$. We can compute the expectation value by

$$\mathbb{E}(\hat{C}_{AR_q}) = \mathbb{E}(q_A q_R) = \sqrt{\frac{\tau_A}{2}} V_M = C_{AR_q}, \quad \text{(C2)}$$

and the variance can be defined as

$$V_{\mathrm{Cov}} := \mathrm{Var}(\hat{C}_{AR_q}) \quad \text{(C3)}$$

with

$$\begin{aligned} \mathrm{Var}(\hat{C}_{AR_q}) &= \frac{1}{m}\mathrm{Var}\,(q_A q_R) \\ &= \frac{1}{2m}\big[\tau_A \mathrm{Var}\big(q_A^2\big) + \tau_B \mathrm{Var}(q_A q_B)\big] \\ &\quad + \mathrm{Var}(q_A q_N), \quad \text{(C4)} \\ &= \frac{1}{m}\left(\tau_A V_M^2 + \frac{\tau_B}{2} V_M^2 + V_M V_{q,N}\right) \\ &= \frac{1}{m}\left(\tau_A + \frac{\tau_B}{2}\right) V_M^2 \left[1 + \frac{V_{q,N}}{\big(\tau_A + \frac{\tau_B}{2}\big) V_M}\right], \quad \text{(C5)} \end{aligned}$$

where we have considered the independence of the variables and the second-order moments of the normal distribution.

Therefore, we can derive the mean and variance for the estimator of τ_A. We rewrite the estimator as

$$\hat{\tau}_{Aq} = \frac{2V_{\mathrm{Cov}}}{V_M^2}\left(\frac{\hat{C}_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2. \quad \text{(C6)}$$

Note that the variable $\big(\hat{C}_{AR_q}/\sqrt{V_{\mathrm{Cov}}}\big)^2$ is χ^2 distributed, i.e.,

$$\left(\frac{\hat{C}_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2 \sim \chi^2\left[1, \left(\frac{C_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2\right], \quad \text{(C7)}$$

with expectation value

$$\mathbb{E}\left[\left(\frac{\hat{C}_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2\right] = 1 + \left(\frac{C_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2 \tag{C8}$$

and variance

$$\mathrm{Var}\left[\left(\frac{\hat{C}_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2\right] = 2\left[1 + 2\left(\frac{C_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2\right]. \tag{C9}$$

The expectation value of $\hat{\tau}_{Aq}$ is then given by

$$\begin{aligned}\mathbb{E}(\hat{\tau}_{Aq}) &= \frac{2V_{\mathrm{Cov}}}{V_M^2}\left[1 + \left(\frac{C_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2\right] \\ &= \frac{2C_{AR_q}^2}{V_M^2} + \mathcal{O}(1/m) = \tau_A + O(1/m)\end{aligned} \tag{C10}$$

and its variance is

$$\begin{aligned}\mathrm{Var}(\hat{\tau}_{Aq}) &= \frac{4V_{\mathrm{Cov}}^2}{V_M^4} 2\left[1 + 2\left(\frac{C_{AR_q}}{\sqrt{V_{\mathrm{Cov}}}}\right)^2\right] \\ &= \frac{16V_{\mathrm{Cov}}C_{AR_q}^2}{V_M^4} + O(1/m^2).\end{aligned} \tag{C11}$$

By replacing Eqs. (C2) and (C5), we obtain

$$\begin{aligned}\mathrm{Var}(\hat{\tau}_{Aq}) &= \frac{16}{mV_M^4}\left(\tau_A + \frac{\tau_B}{2}\right)\frac{V_M^4\tau_A}{2}\left[1 + \frac{V_{q,N}}{\left(\tau_A + \frac{\tau_B}{2}\right)V_M}\right] \\ &\quad + O(1/m^2),\end{aligned} \tag{C12}$$

$$\begin{aligned}&= \frac{8\tau_A}{m}\left(\tau_A + \frac{\tau_B}{2}\right)\left[1 + \frac{V_{q,N}}{\left(\tau_A + \frac{\tau_B}{2}\right)V_M}\right] \\ &\quad + O(1/m^2).\end{aligned} \tag{C13}$$

Clearly, as previously noted in Ref. [60], the estimator of the transmissivity $\hat{\tau}_{Aq}$ is only asymptotically unbiased. In fact, the standard deviation $\sqrt{\mathrm{Var}(\hat{\tau}_{Aq})}$ is of order $1/\sqrt{m}$ while the bias goes as $1/m$. As we consider $m > 10^5$ in our analysis, the value of the bias become rapidly negligible as $m \gg 1$, and the use of estimators $\hat{\tau}_A$ and $\hat{\tau}_B$ are very accurate.

2. Variance of the excess noise

Also in our estimation procedure for the MDI protocol, in analogy to the theory developed in Ref. [60], the variance of the estimator can be obtained replacing the estimator of τ_A (τ_B) with its value. This simplifies the calculations. Now, we can assume that any uncertainty in the estimator of the excess noise obtained from broadcast results of relay's measurements on quadrature q

$$\hat{V}_{q,\epsilon} = \frac{1}{m}\sum_{i=1}^{m}\left[R_{q,i} - \frac{\sqrt{\hat{\tau}_B}B_{q,i} - \sqrt{\hat{\tau}_A}A_{q,i}}{\sqrt{2}}\right]^2 - 1 \tag{C14}$$

comes only from variables $R_{q,i}$, $A_{q,i}$, and $B_{q,i}$. We then have that the expression inside square brackets is normally distributed with zero mean and variance $V_{q,N}$. In addition to this, one also has that the following expression

$$Y := \sum_{i=1}^{m}\left(\frac{R_{q,i} - \left(\sqrt{\tau_B}B_{q,i} - \sqrt{\tau_A}A_{q,i}\right)/\sqrt{2}}{\sqrt{V_{q,N}}}\right)^2 \sim \chi^2(m,0) \tag{C15}$$

is χ^2 distributed and has mean $\mathbb{E}(Y) = m$ and variance $\mathrm{Var}(Y) = 2m$. This allows us to approximate the sum of Eq. (C14) with $V_{q,N}Y$ when we assume large values for m, obtaining the expectation value

$$\mathbb{E}(\hat{V}_{q,\epsilon}) \approx \frac{1}{m}V_{q,N}\mathbb{E}(Y) - 1 = V_{q,\epsilon} \tag{C16}$$

and the variance

$$\mathrm{Var}(\hat{V}_{q,\epsilon}) \approx \frac{1}{m^2}V_{q,N}^2\mathrm{Var}(Y) = \frac{2}{m}V_{q,N}^2. \tag{C17}$$

[1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).
[2] B. Schneier, *Applied Cryptography* (John Wiley & Sons, New York, 1996).
[3] H. J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **81**, 5932 (1998).
[4] L. M. Duan, M. Lukin, J. I. Cirac, and P. Zoller, Nature (London) **414**, 413 (2001).
[5] H. J. Kimble, Nature (London) **453**, 1023 (2008).
[6] S. Pirandola and S. L. Braunstein, Nature (London) **532**, 169 (2016).
[7] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, Nat. Commun. **8**, 15043 (2017); see also arXiv:1510.08863 and arXiv:1512.04945.
[8] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).
[9] V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998).
[10] V. Vedral, Rev. Mod. Phys. **74**, 197 (2002).
[11] S. Pirandola, arXiv:1601.00966.
[12] R. Laurenza and S. Pirandola, Phys. Rev. A **96**, 032318 (2017).
[13] T. P. W. Cope, L. Hetzel, L. Banchi, and S. Pirandola, Phys. Rev. A **96**, 022323 (2017).
[14] R. Laurenza, S. L. Braunstein, and S. Pirandola, arXiv:1706.06065.
[15] S. Pirandola and C. Lupo, Phys. Rev. Lett. **118**, 100502 (2017).
[16] R. Namiki, L. Jiang, J. Kim, and N. Lütkenhaus, Phys. Rev. A **94**, 052304 (2016).
[17] K. Bradler, T. Kalajdzievski, G. Siopsis, and C. Weedbrook, arXiv:1607.05916.
[18] B. A. Bash, N. Chandrasekaran, J. H. Shapiro, and S. Guha, arXiv:1604.08582.
[19] J. H. Shapiro, Q. Zhuang, Z. Zhang, J. Dove, and F. N. Wong, in Lasers Congress 2016 (ASSL, LSC, LAC), OSA Technical Digest (online), paper LTu5B.1 (unpublished).
[20] C. Ottaviani, R. Laurenza, T. P. W. Cope, G. Spedalieri, S. L. Braunstein, and S. Pirandola, Proc. SPIE **9996**, 999609 (2016).

[21] M. Pant, H. Krovi, D. Englund, and S. Guha, Phys. Rev. A **95**, 012304 (2017).
[22] F. Ewert and P. van Loock, Phys. Rev. A **95**, 012327 (2017).
[23] A. Khalique and B. C. Sanders, Opt. Eng. **56**, 016114 (2017).
[24] F. Rozpedek, K. Goodenough, J. Ribeiro, N. Kalb, V. Caprara Vivoli, A. Reiserer, R. Hanson, S. Wehner, and D. Elkouss, arXiv:1705.00043.
[25] K. Bradler and C. Weedbrook, arXiv:1709.01758.
[26] N. Lo Piparo, N. Sinclair, and M. Razavi, arXiv:1707.07814.
[27] N. Lo Piparo and M. Razavi, in Conference on Lasers and Electro-Optics (CLEO), 5–10 June 2016 (unpublished).
[28] N. Lo Piparo, M. Razavi, and W. J. Munro, arXiv:1708.06532.
[29] M. Pant, H. Krovi, D. Towsley, L. Tassiulas, L. Jiang, P. Basu, D. Englund, and S. Guha, arXiv:1708.07142.
[30] S. L. Braunstein and P. van Loock, Rev. Mod. Phys. **77**, 513 (2005).
[31] C. Weedbrook, S. Pirandola, R. Garcí a-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Rev. Mod. Phys. **84**, 621 (2012).
[32] E. Diamanti and A. Leverrier, Entropy **17**, 6072 (2015).
[33] F. Grosshans and P. Grangier, Phys. Rev. Lett. **88**, 057902 (2002).
[34] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, Phys. Rev. Lett. **93**, 170504 (2004).
[35] V. C. Usenko and F. Grosshans, Phys. Rev. A **92**, 062337 (2015).
[36] S. Pirandola, S. Mancini, S. Lloyd, and S. L. Braunstein, Nat. Phys. **4**, 726 (2008).
[37] T. Gehring, C. S. Jacobsen, and U. L. Andersen, Quantum Inf. Comput. **16**, 1081 (2016).
[38] F. Grosshans, G. Van Ache, J. Wenger, R. Brouri, N. J. Cerf, and P. Grangier, Nature (London) **421**, 238 (2003).
[39] J. Lodewyck, M. Bloch, R. García-Patrón, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N. J. Cerf, R. Tualle-Brouri, S. W. McLaughlin *et al.*, Phys. Rev. A **76**, 042305 (2007).
[40] L. S. Madsen, V. C. Usenko, M. Lassen, R. Filip, and U. L. Andersen, Nat. Commun. **3**, 1083 (2012).
[41] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti, Nat. Photon. **7**, 378 (2013).
[42] D. Huang, P. Huang, D. Lin, and G. Zeng, Sci. Rep. **6**, 19201 (2016).
[43] R. Filip, Phys. Rev. A **77**, 022310 (2008).
[44] C. Weedbrook, S. Pirandola, S. Lloyd, and T. C. Ralph, Phys. Rev. Lett. **105**, 110501 (2010).
[45] V. C. Usenko and R. Filip, Phys. Rev. A **81**, 022318 (2010).
[46] C. Weedbrook, S. Pirandola, and T. C. Ralph, Phys. Rev. A **86**, 022318 (2012).
[47] C. Weedbrook, C. Ottaviani, and S. Pirandola, Phys. Rev. A **89**, 012309 (2014).
[48] V. C. Usenko and R. Filip, Entropy **18**, 20 (2016).
[49] C. S. Jacobsen, T. Gehring, and U. L. Andersen, Entropy **17**, 4654 (2015).
[50] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Ghering, C. S. Jacobsen, and U. L. Andersen, Nat. Photon. **9**, 397 (2015).
[51] C. Ottaviani, G. Spedalieri, S. L. Braunstein, and S. Pirandola, Phys. Rev. A **91**, 022320 (2015).
[52] S. L. Braunstein and S. Pirandola, Phys. Rev. Lett. **108**, 130502 (2012).
[53] H. K. Lo, M. Curty, and B. Qi, Phys. Rev. Lett. **108**, 130503 (2012).
[54] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).
[55] F. Furrer, T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner, Phys. Rev. Lett. **109**, 100502 (2012).
[56] A. Leverrier, Phys. Rev. Lett. **114**, 070501 (2015).
[57] A. Leverrier, Phys. Rev. Lett. **118**, 200501 (2017).
[58] V. Scarani and R. Renner, Phys. Rev. Lett. **100**, 200501 (2008).
[59] A. Leverrier, F. Grosshans, and P. Grangier, Phys. Rev. A **81**, 062343 (2010).
[60] L. Ruppert, V. C. Usenko, and R. Filip, Phys. Rev. A **90**, 062310 (2014).
[61] I. Devetak and A. Winter, Proc. R. Soc. London, Ser. A **461**, 207 (2005).
[62] G. Spedalieri, C. Ottaviani, S. L. Braunstein, T. Gehring, C. S. Jacobsen, U. L. Andersen, and S. Pirandola, Proc. SPIE **9648**, 96480Z (2015).
[63] C. Ottaviani, S. Mancini, and S. Pirandola, Phys. Rev. A **95**, 052310 (2017).
[64] C. Ottaviani and S. Pirandola, Sci. Rep. **6**, 22225 (2016).
[65] C. Ottaviani, S. Mancini, and S. Pirandola, Phys. Rev. A **92**, 062323 (2015).
[66] C. Ottaviani, G. Spedalieri, S. L. Braunstein, and S. Pirandola, arXiv:1509.04144.
[67] G. Spedalieri, C. Ottaviani, and S. Pirandola, Open Syst. Infor. Dynam. **20**, 1350011 (2013).
[68] S. Pirandola, S. L. Braunstein, and S. Lloyd, Phys. Rev. Lett. **101**, 200504 (2008).
[69] S. Pirandola, New J. Phys. **15**, 113046 (2013).
[70] P. Jouguet, S. Kunz-Jacques, and A. Leverrier, Phys. Rev. A **84**, 062317 (2011).
[71] M. Milicevic, C. Feng, L. M. Zhang, and P. G. Gulak, arXiv:1702.07740.
[72] X. Wang, Y.-C. Zhang, Z. Li, B. Xu, S. Yu, and H. Guo, arXiv:1703.04916.
[73] C. Lupo, C. Ottaviani, P. Papanastasiou, and S. Pirandola, arXiv:1704.07924.
[74] X. Zhang, Y.-C. Zhang, Y. Zhao, X. Wang, S. Yu, and H. Guo, arXiv:1707.05931.