HAL
archives-ouvertes.fr

Continuous-variable quantum-key-distribution protocols with a non-Gaussian modulation

Anthony Leverrier, Philippe Grangier

▶ To cite this version:

Anthony Leverrier, Philippe Grangier. Continuous-variable quantum-key-distribution protocols with a non-Gaussian modulation. Physical Review A, American Physical Society, 2011, 83 (4), pp.042312. 10.1103/PhysRevA.83.042312 . hal-00624775

HAL Id: hal-00624775
https://hal-iogs.archives-ouvertes.fr/hal-00624775

Submitted on 1 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Continuous-variable quantum-key-distribution protocols with a non-Gaussian modulation

Anthony Leverrier
ICFO-Institut de Ciencès Fotòniques, F-08860 Castelldefels (Barcelona), Spain

Philippe Grangier
Laboratoire Charles Fabry, Institut d'Optique, CNRS, Univ. Paris-Sud, Campus Polytechnique, RD 128, F-91127 Palaiseau Cedex, France
(Received 15 January 2011; published 11 April 2011)

In this paper, we consider continuous-variable quantum-key-distribution (QKD) protocols which use non-Gaussian modulations. These specific modulation schemes are compatible with very efficient error-correction procedures, hence allowing the protocols to outperform previous protocols in terms of achievable range. In their simplest implementation, these protocols are secure for any linear quantum channels (hence against Gaussian attacks). We also show how the use of decoy states makes the protocols secure against arbitrary collective attacks, which implies their unconditional security in the asymptotic limit.

DOI: 10.1103/PhysRevA.83.042312 PACS number(s): 03.67.Dd, 42.50.−p

I. INTRODUCTION

The first potentially real-life application of the emerging field of quantum information is arguably quantum-key-distribution (QKD), which allows two distant parties to establish a secret key over an *a priori* unsecure communication channel [1]. The importance of a QKD protocol is usually measured through three main criteria: its *practicality*, its *performance*, and its *security*.

Among all existing protocols, those based on continuous variables appear quite appealing from a *practical* point of view [2]. For instance, the Grosshans-Grangier 2002 (GG02) [3] protocol simply requires the preparation of coherent states and their detection via homodyne (or heterodyne) detection. Moreover, the *security* of this protocol is established against collective attacks [4–6] which are optimal in the asymptotic limit [7]. There exist different ways to quantify the *performance* of a given protocol, but the usual figure of merit is the secret key rate that can be achieved as a function of the distance. Continuous-variable protocols such as the GG02 protocol perform quite well for short distances, but seem unfortunately limited to much shorter distances than their discrete-variable counterparts (see Fig. 4 of Ref. [1] for a comparison of the performance of various protocols). In particular, while discrete-variable protocols allow one to distribute secret keys over distances larger than 100 km, the GG02 protocol has only been implemented for a distance of 25 km [8,9] and there is not much hope of increasing its range well beyond 50 km [10].

In this paper, we introduce continuous-variable (CV) QKD protocols which share the practicality and security of the GG02 protocol, but greatly improve its performance, especially in terms of achievable distance. In Sec. II, we quickly review the status of current CV protocols. In Sec. III, we present the specific modulation schemes of our protocols. We then introduce the concept of *decoy states* for CV QKD protocols in Sec. IV. In Sec. V, we prove the security of our protocols against collective attacks, and we finally discuss their performance in Sec. VI.

II. LIMITATIONS OF CURRENT CV PROTOCOLS

Proving the security of a QKD protocol is usually a difficult task, but the situation is even worse for continuous-variable QKD protocols, because the relevant Hilbert space is infinite dimensional. The task is difficult for (at least) one specific reason: the bipartite state ρ_{AB} shared by Alice and Bob (in the entanglement-based version of the protocol) has an infinite number of degrees of freedom. This means that a full tomography of this state is hopeless, even in the case of collective attacks, where Alice and Bob share many copies of the same state.

Fortunately, one can take advantage of extremality properties of Gaussian states [11] to show that the eavesdropper's information is upper-bounded by the information she would have if Alice and Bob shared instead the state ρ_{AB}^{G}, the Gaussian state with the same first two moments as ρ_{AB} [4–6]. Hence, knowing the first two moments of the state ρ_{AB}, that is, a *finite* number of parameters, is sufficient to bound Eve's information. Furthermore, taking advantage of specific symmetries of the protocols in phase space [12], it is possible to reduce this number to only 3, namely, the variances of Alice and Bob's reduced states and the covariance. That so few parameters are indeed sufficient to bound Eve's information is quite remarkable, but it should be noted that this is also a necessity if one wants to take finite-size effects into account. Indeed, it was shown in [13] that estimating these parameters with a precision compatible with a secret key rate is already very challenging in terms of resources.

Because of the constraints imposed by finite-size effects, the only theoretical tool presently available to prove the security of a continuous-variable QKD protocol is therefore this Gaussian optimality. This technique unfortunately comes at a price, namely, that if the true quantum state ρ_{AB} is not sufficiently close to a Gaussian state, then the bound on Eve's information will not be tight enough to still get secret bits at the end of the protocol. By construction, this bound is indeed only tight for Gaussian states, and it turns out that it degrades very rapidly as the non-Gaussianity of the state (and consequently of the protocol) increases.

 ©2011 American Physical Society

This observation leads us to the unavoidable conclusion that with the theoretical techniques available today, protocols using a Gaussian modulation (i.e., the GG02 protocol, possibly with a heterodyne detection [14]) are optimal among all continuous-variable QKD protocols, and using any other modulation scheme can only lead to worse performances.[1]

A natural question then arises: why should one consider new protocols involving specific non-Gaussian modulations since *theoretically*, they cannot beat a Gaussian modulation? The reason is that, *in practice*, protocols with a Gaussian modulation do not perform as well as expected, especially with regards to long-distance communication. This is due to the fact that error correction is very hard to implement for a Gaussian modulation, thereby making the effective secret key rate drop to zero at distances of the order of 50 km, even if the theoretical key rate is strictly positive.

The protocols we introduce in the present paper outperform previous protocols in terms of maximal range. This is achieved thanks to a combination of two facts: first, the specific modulation schemes allow one to extract information much more efficiently than with a Gaussian modulation; second, there is a regime (for the variance of the modulation) where the protocols are still significantly close to a Gaussian protocol, hence making the bound on Eve's information tight enough to be useful.

III. ALTERNATIVE MODULATION SCHEMES

The main argument for switching from a Gaussian modulation to a non-Gaussian one is *not* because it might make the eavesdropping more difficult (available theoretical tools are not powerful enough to answer this question). The reason is that present coding techniques do not allow for a very efficient reconciliation procedure at (very) low signal-to-noise ratio (SNR) when Alice's data are Gaussian. Remember that a QKD protocol typically consists of three phases: first, Alice and Bob exchange quantum signals, perform measurements, and obtain correlated classical data, say, vectors $\mathbf{x}$ and $\mathbf{y}$; second, in the *reconciliation* step, they use classical error correction techniques to agree on a common (errorless) bit string $\mathbf{u}$, and finally, they apply *privacy amplification* to obtain a secure key from $\mathbf{u}$.

The importance of the reconciliation phase is specific to continuous-variable protocols for two reasons. First, one has to deal with continuous values for *both* Alice and Bob's variables instead of bits, which is rather unusual in the field of digital communication. Indeed, even when analog signals are used to transmit information on noisy (classical) channels, the modulation is almost always discrete and not continuous. The second reason is that contrary to discrete-variable QKD protocols where the error rate is always below some small constant (to guarantee the security of the protocol), the SNR can be arbitrarily small for continuous-variable protocols (and is actually very small as one tries to increase the range of the protocol[2]).

For continuous-variable QKD protocols, information is encoded in phase space, in general in the quadratures of coherent states.[3] More precisely, if the classical information she wants to send is described by a vector $\mathbf{x} = (x_1, x_2, \ldots, x_{2n})$, then Alice prepares the n coherent states $|x_1 + ix_2\rangle, \ldots, |x_{2n-1} + ix_{2n}\rangle$ and sends them through the quantum channel.

There are two possibilities concerning the detection: Bob can perform either homodyne or heterodyne measurements. In the case of a homodyne detection, for each state, Bob chooses randomly which quadrature to measure. He then obtains an n-dimensional classical vector $\mathbf{y}$ and later informs Alice of his choices of quadratures. In the case of a heterodyne detection, Bob ends up with a $2n$-dimensional vector $\mathbf{y}$. Hence, Alice and Bob share twice as many data for the heterodyne protocol than for the homodyne one. However, a heterodyne detection adds 3 dB of noise to the data, and in practice the performances of both schemes are quite similar.

In order to complete the key distribution, two additional steps are required. First Alice and Bob proceed with the reconciliation of their classical data in order to agree on a common bit string u. Here, we restrict ourselves to a *reverse* reconciliation [15], meaning that only Bob can send classical information to Alice (in contrast to *direct* reconciliation where Alice would send some side information to help Bob correct his errors). Second, they use privacy amplification in order to obtain a secret key from u: this can only be done once they have an upper bound on Eve's information about u.

In this paper, we consider protocols for which the reconciliation can be performed efficiently (which is not the case of the protocols using a Gaussian modulation), and for which Eve's information can be bounded if she is restricted to collective attacks. Security against general attacks ("unconditional security") is then obtained in the asymptotic limit, thanks to a de Finetti representation theorem for infinite-dimensional quantum systems [7].

To be more specific, the protocols considered here are characterized as the ones for which the *reverse* reconciliation problem can be reduced to the channel coding problem for the binary-input additive white Gaussian noise (BI-AWGN) channel, which can itself be tackled with standard techniques (efficient error-correcting codes). One such protocol is the four-state protocol considered in [16], but it turns out that other, more efficient, *continuous* modulation schemes are also possible.

[1] Note that restricting the eavesdropper to specific classes of attacks, such as beam-splitter attacks, allows one to compute tighter bounds on the secret key rate, even for non-Gaussian modulations, but this is not the case if one is not willing to arbitrarily restrict the possible attacks of the eavesdropper.

[2] As a first-order estimate, if the variance of modulation is V_A and the transmission of the channel is $T \approx 10^{-0.02d}$ (in fiber if d is the channel length in kilometers), then the SNR is roughly given by TV_A. The modulation variance on Alice's side, V_A is a free parameter of the protocol which can be optimized for a quantum channel, but numerical simulations show that the optimal value for V_A practically does not depend on the transmission T of the channel. For this reason, the SNR is roughly proportional to the transmission of the channel, and is consequently very small for long distances.

[3] In principle, Alice could alternatively encode information on squeezed states, but this is not practical and does not lead to vastly superior performances.

A. Homodyne detection: Four-state modulation

Let us consider protocols involving a homodyne detection. In this case, the protocols of interest display a discrete modulation with either two or four states, and a basic problem is how to evaluate the transmission channel, and thereby Eve's information. Such a problem does not arise for Gaussian modulations, because the variances and covariances directly measured by Alice and Bob give a covariance matrix, which is all that is needed to characterize the worst possible attack by Eve, which is a Gaussian attack according to the Gaussian optimality theorem. On the other hand, for a non-Gaussian modulation, the noises and correlations measured by Alice and Bob cannot be directly connected to a relevant covariance matrix, and the Gaussian optimality theorem does not apply directly.

Among possible approaches, two-state protocols were studied in [17], and the authors proved the security of the protocols against any collective attacks, but with the *caveat* that a complete tomography of the state was required. In [16,18], the authors considered the noises and correlations measured by Alice and Bob from their non-Gaussian modulation, and then considered the Gaussian attacks associated with these values as optimal. However, this approach has the implicit assumption that the transmission channel can be considered as *linear*, which means that it is intrinsically characterized by a transmission T and an excess noise ξ (see details in Appendix A). Gaussian channels are obvious examples of linear channels, but a linear channel may also be non-Gaussian. Since the proofs of [16,18] only involve estimating the second moments of Alice and Bob's correlations, they are compatible with a practical implementation taking into account finite-size effects [13], but they are not fully general.

In the present paper, we extend the security proof of discrete modulation protocols against arbitrary collective attacks. Our proof still only requires us to estimate two quantities: the transmission T and excess noise ξ of the quantum channel, without carrying out a full channel tomography. But achieving this estimation requires the use of decoy states, as will be explained in Sec. IV.

Since the four-state protocol always outperforms the two-state protocol, we will only discuss the case of the former in the present paper. In this protocol, the modulation consists of four coherent states: $|\alpha e^{i\pi/4}\rangle$, $|\alpha e^{3i\pi/4}\rangle$, $|\alpha e^{-3i\pi/4}\rangle$, and $|\alpha e^{-i\pi/4}\rangle$, where α is a positive number. The modulation variance V_A is given by $V_A = 2\alpha^2$.

The practical implementation of the reverse reconciliation problem for the four-state protocol is discussed in detail in Appendix B.

B. Heterodyne detection: Non-Gaussian continuous modulations

Before introducing the modulation schemes compatible with a heterodyne detection, let us say a few words concerning the reconciliation procedure. The main difficulty here lies in the fact that we need a reverse reconciliation. Indeed, the side information sent by Bob must help Alice without giving Eve any relevant information. The only schemes where side information seems to have these properties (while being efficiently computable by Bob) are when it describes rotations in particular dimensions, namely, dimensions 1, 2, 4, and 8 [10]. This surprising result is a consequence of the fact that the only real division algebras are the real numbers ($\mathbb{R}$), the complex numbers ($\mathbb{C} \cong \mathbb{R}^2$), the quaternions ($\mathbb{H} \cong \mathbb{R}^4$), and the octonions ($\mathbb{O} \cong \mathbb{R}^8$). Indeed, one can show that the possibility of an efficient reconciliation protocol (in terms of computation complexity as well as classical communication) is intimately connected with the existence of a division operation for the data considered. In particular, identifying vectors in $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^4$, or $\mathbb{R}^8$ with units of the real numbers, the complex numbers, the quaternions, or the octonions allows one to take advantage of the division structure naturally associated with these ensembles. For instance, the reconciliation protocol of the four-state protocol exploits this property in dimension 1. The modulation schemes we consider now exploit it for dimensions 2, 4, and 8.

First, note that in the case of the four-state protocol, Alice chooses the value of each quadrature uniformly on the (zero-dimensional) sphere $\mathcal{S}^0 = \{-1,1\}$ in dimension 1. (This value is then appropriately rescaled to obtain the desired variance of modulation.) For this reason, we will sometimes refer to the four-state protocol as the one-dimensional protocol in the rest of the paper.

The modulation schemes we consider now are simply the generalizations to dimensions 2, 4, and 8. Hence, in the d-dimensional protocol (with $d \in \{2,4,8\}$), Alice draws random variables uniformly on the sphere $\mathcal{S}^{d-1}$ in dimension d.

For $d = 2$, Alice draws n points on the unit circle: $\{(x_1^1,x_1^2),(x_2^1,x_2^2),\ldots,(x_n^1,x_n^2)\}$, and sends the n coherent states $|x_1^1 + ix_1^2\rangle,\ldots,|x_n^1 + ix_n^2\rangle$, where the variables are rescaled by a factor of α.

For $d = 4$, Alice draws $n/2$ points on the sphere $\mathcal{S}^3$: $\{(x_1^1,x_1^2,x_1^3,x_1^4),\ldots,(x_{n/2}^1,x_{n/2}^2,x_{n/2}^3,x_{n/2}^4)\}$, and sends the n coherent states $|x_1^1 + ix_1^2\rangle,|x_1^3 + ix_1^4\rangle,\ldots,|x_{n/2}^3 + ix_{n/2}^4\rangle$, where the variables are rescaled by a factor of $\alpha\sqrt{2}$.

For $d = 8$, Alice draws $n/4$ points on the sphere $\mathcal{S}^7$: $\{(x_1^1,x_1^2,x_1^3,x_1^4,x_1^5,x_1^6,x_1^7,x_1^8),\ldots,(x_{n/4}^1,x_{n/4}^2,x_{n/4}^3,x_{n/4}^4,x_{n/4}^5,x_{n/4}^6,x_{n/4}^7,x_{n/4}^8)\}$, and sends the n coherent states $|x_1^1 + ix_1^2\rangle,|x_1^3 + ix_1^4\rangle,\ldots,|x_{n/4}^7 + ix_{n/4}^8\rangle$, where the variables are rescaled by a factor of 2α.

The procedure to reduce the reverse reconciliation problem in these three scenarios to the usual problem of channel coding for the BI-AWGN channel is explained in detail in Appendix B.

Let us say a few words about what we mean by *efficient* reconciliation procedure in the context of QKD. Usually in the field of computer science, an algorithm is said to be efficient if the amount of resources (e.g., running time, randomness generation, classical communication, etc.) it consumes grow at most polynomially with the natural size of the problem. In the case of a reconciliation procedure, the problem size is given by the length n of the vectors Alice and Bob try to agree on. As n usually takes very large values (for instance, 10^{10} or 10^{12}), an algorithm requiring resources scaling as n^2 or n^3 is obviously unacceptable: only a linear scaling is compatible with a practical implementation. Moreover, the factor of proportionality should be small enough. For $d = 1, 2, 4,$ or 8, the reconciliation procedure introduced in [10] requires that Bob draws one random bit and transmits classically one

real number to Alice per exchanged signal. The computational complexity of the protocol (not including the decoding of the error-correcting code) is linear with n. For values of d strictly greater than 8, the naive approach (which consists in drawing random transformations uniformly in the orthogonal group O_d and transmitting them to Alice) requires that Bob draw d random variables from a normal distribution (instead of only 1 bit) and send d real values to Alice for each exchanged quantum signal, which is prohibitive for a realistic implementation.

It should be emphasized that the higher the dimension, the higher the secret key rate of the QKD protocol. The reason for this is the very specific technique that we use to bound Eve's information. Our bound depends only on the covariance matrix of Alice and Bob's bipartite state in the entanglement-based version of the protocol, and therefore is tight only when the state is Gaussian. Fortunately, if the state is almost Gaussian, then the bound is good enough for our purpose. Because of this, we want to use a protocol as Gaussian as possible. It turns out that considering modulations in higher and higher dimensions brings us closer and closer to the Gaussian modulation for which the bound is tight. Indeed, a Gaussian modulation of variance 1 can be seen as drawing uniformly a random point of the sphere of radius $\sqrt{d}$ in $\mathbb{R}^d$ as d tends to infinity. Hence, the GG02 protocol with a Gaussian modulation can be seen as the d-dimensional protocol with $d = \infty$. Unfortunately, for $d = \infty$, efficient reconciliation techniques at low SNR are not known.

IV. DECOY STATES

As we already mentioned, it is crucial that the security proof of CV QKD protocols requires the estimation of only a few parameters. Ideally, one would like a GG02-type security proof where only the transmission T and the excess noise ξ of the quantum channel need to be estimated. As we will discuss in Sec. V, the security proof of our protocols indeed relies on the fact that one can estimate the covariance matrix of Alice and Bob's bipartite state in the entanglement-based version of the protocol.

A difficulty lies in the fact that Alice and Bob do not perform the entanglement-based version of the protocol (in which case the covariance matrix is directly accessible in the experiment) but use instead the *prepare-and-measure* version. Hence, if Alice encodes the variable x in the quadrature of a state and Bob obtains the result y when measuring this quadrature, they can estimate the three following moments of order 2: Alice's variance $\langle x^2 \rangle$, Bob's variance $\langle y^2 \rangle$, and the covariance $\langle xy \rangle$. Whereas Alice and Bob's variance in the prepare-and-measure scenario are directly related to the respective variances in the entanglement-based scenario, the same is not true for the covariances.

There are two cases where the covariance matrix in the prepare-and-measure protocol allows one to recover the covariance matrix of the state in the entanglement-based scenario, namely, when the quantum channel is linear (see Appendix A), for instance, a Gaussian channel, and when the modulation is Gaussian.

In Refs. [18,19], the security of the protocols considered in the present paper was established in the case of linear channels. Here, we wish to get rid of this hypothesis (which can never be perfectly verified in practice with a finite number of samples), and for this reason, it is necessary to use a Gaussian modulation for the parameter estimation procedure. Unfortunately, it is not *a priori* possible to use two different modulations for key distribution and parameter estimation, because an eavesdropper could use a different strategy in each case. The solution is to add a *third* modulation consisting of decoy states. Let us call "key," "decoy," and "G" the modulations corresponding, respectively, to states used for the key distillation, decoy states, and states used for parameter estimation purposes (a Gaussian, in fact, thermal, distribution). One can define the three following states:

$$\sigma^d_{\text{key}} = \int p_{\text{key}}(\alpha)\, |\alpha\rangle\langle\alpha|\, d\alpha, \tag{1}$$

$$\sigma^d_{\text{decoy}} = \int p_{\text{decoy}}(\alpha)\, |\alpha\rangle\langle\alpha|\, d\alpha, \tag{2}$$

$$\sigma^d_{\text{G}} = \int p_{\text{G}}(\alpha)\, |\alpha\rangle\langle\alpha|\, d\alpha, \tag{3}$$

where $\alpha \in \mathbb{R}^d$ for the d-dimensional protocol. (In particular, $|\alpha\rangle$ refers here to $d/2$ coherent states.) In these expressions, the probability distribution p_{key} is the uniform measure of the sphere $\mathcal{S}^{d-1}$ (with radius $\alpha\sqrt{d/2}$) and p_{G} is the Gaussian distribution $\mathcal{N}(0,\alpha^2/2)^{\otimes d}$ in d dimensions. In other words, p_{key} corresponds to the modulation schemes described in Sec. III, and p_{G} is the Gaussian distribution of the GG02 protocol.

If the probability distribution p_{decoy} is chosen such that

$$p\,\sigma^d_{\text{key}} + (1-p)\,\sigma^d_{\text{decoy}} = \sigma^d_{\text{G}}, \tag{4}$$

where p is a weight between 0 and 1, then the state sent by Alice to use for parameter estimation is indistinguishable from that used to distill a key (or as a decoy). The idea is that after the exchange of quantum states is complete, Alice can announce to Bob which states he can use for the key, which states he can discard (decoys), and which states should be used for parameter estimation.

If, in principle, the form of σ^d_{decoy} does not matter (it could be any state, not necessarily of the form 2), this is no longer true if we require the protocol to be practical. Indeed, for this reason, we impose the extra constraint that σ^d_{decoy} should be obtained as a mixture of coherent states, i.e., be of the form 2. We discuss this in detail in Appendix C, where we describe two techniques for finding the appropriate decoy states.

To summarize, the modulation that is used in the protocols is a mixture of three particular modulations. Let us note p_{est} the fraction of states that Alice and Bob want to use for parameter estimation purposes (this fraction can be optimized numerically but in a typical scenario, its value can be around 50%). Then for each state she sends, Alice will choose either the modulation $p_{\text{key}}(\alpha)$ with probability $p(1-p_{\text{est}})$ or modulation $p_{\text{G}}(\alpha)$ with probability p_{est} or send a decoy state with probability $(1-p)(1-p_{\text{est}})$.

V. SECURITY AGAINST COLLECTIVE ATTACKS

In this paper, we restrict ourselves to the case of collective attacks since they are optimal in the asymptotic limit [7]. The (asymptotic) secret key rate K is then given by [20]

$$K = \beta I(A;B) - \chi(B;E), \tag{5}$$

where β is the reconciliation efficiency, $I(A;B)$ is the classical mutual information between Alice and Bob's data [for the data corresponding to the modulation $p_{\text{key}}(\alpha)$], and $\chi(B;E)$ is the Holevo quantity

$$\chi(B;E) = S(\rho_E) - \sum_y p(y) S(\rho_{E|y}), \tag{6}$$

where S is the von Neumann entropy, y is Bob's measurement result obtained with probability $p(y)$, $\rho_{E|y}$ is the corresponding state of Eve's ancilla, and $\rho_E = \sum_y p(y)\rho_{E|y}$ is Eve's partial state.

Note that this rate should be modified to take finite-size effects into account. For simplicity, we only consider the asymptotic rate here, but a complete analysis of finite-size effects can be found in Ref. [13].

Since the quantity $\beta I(A;B)$ is directly observable in practice, the goal of the security proof consists in deriving an upper bound for the quantity $\chi(B;E)$ which should be a function of parameters accessible in an experiment. In our case, we will find a bound which only depends on three parameters: the variance of modulation V_A which is chosen by Alice, the transmission T, and the excess noise ξ of the quantum channel which can be estimated with the technique described in Ref. [13].

We now consider the entanglement-based version of our protocols. In this scenario, Alice prepares n two-mode squeezed vacuum states, keeps one half of each state and sends the second half to Bob through the quantum channel.

Let us introduce some notation. In the following, we will consider bipartite states, either before or after the quantum channel. We use the superscript 0 to denote states before the quantum channel. Moreover, the action of the quantum channel can be described by a map $\mathbb{1} \otimes \mathcal{T}$ where the identity acts on the first part of the state (namely, Alice's state) and the quantum channel $\mathcal{T}$ acts nontrivially only on the second part of the state.

The three states of interest are ρ_{G}^0, ρ_{key}^0, and ρ_{decoy}^0, which are the Schmidt purifications of the states σ_{G}^d, σ_{key}^d, and σ_{decoy}^d, respectively. (Note, for instance, that ρ_{G}^0 is a two-mode squeezed vacuum: $\rho_{\text{G}}^0 = |\text{EPR}\rangle\langle\text{EPR}|$.) After the quantum channel, these three states become, respectively, ρ_{G}, ρ_{key}, and ρ_{decoy}.

The main idea of the security proof is that one can bound Eve's information by a function of the covariance matrix of the state used to distill the key: ρ_{key}. In the protocol, Alice always starts with a two-mode squeezed state ρ_{G}^0, but she can choose between two measurement strategies depending on whether a given state will be used for key distillation or for parameter estimation.

We also introduce a general measurement acting on Alice's part $\{\Pi_d, \mathbb{1} - \Pi_d\}$ such that, when applied to the two-mode squeezed vacuum ρ_{G}^0, the result corresponding to Π_d prepares the state ρ_{key}^0 used in the d-dimensional protocol while the second result prepares ρ_{decoy}^0 (see Appendix D for a description of this measurement).

For each state, Alice chooses randomly between key distillation and parameter estimation. The fraction of each task should be optimized taking into account all finite-size effects. If a state is dedicated to parameter estimation, Alice simply performs a heterodyne detection on her part and Bob proceeds as usual. At the end of the protocol, Alice and Bob can compare their statistics and compute the covariance matrix Γ_{G} of the state ρ_{G}. For this, we do not need to make any assumption (for instance, of linearity) concerning the quantum channel.

If a state is to be used for key distillation, then Alice performs the generalized measurement $\{\Pi_d, \mathbb{1} - \Pi_d\}$ on her half of the state, thus preparing ρ_{key}^0 with probability p and ρ_{decoy}^0 with probability $1 - p$. States corresponding to the result $\mathbb{1} - \Pi_d$ will later be discarded. Finally, only the states ρ_{key}^0 are used for key distillation. Let us note Γ_{key}^0 (Γ_{key}) the covariance matrix of these states before (after) the quantum channel.

It was proven in [4] that the quantity $\chi(B;E)$ can always be upper bounded by a function of the covariance matrix of Alice and Bob's bipartite state. Here, the state used for key distillation is ρ_{key}, meaning that one can bound $\chi(B;E)$ with a (known) function of Γ_{key} (see Ref. [8] for the precise form of this function).

All that is left to do is therefore to compute the covariance matrix Γ_{key}. Note that the covariance matrix Γ_{key}^0 of the state before the quantum channel can be computed and only depends on the modulation variance (as well as the dimension of the modulation scheme). Details on how to compute this covariance matrix are given in Appendix E. The covariance matrix Γ_{key}^0 has the following form (with the convention x_A, p_A, x_B, p_B):

$$\Gamma_{\text{key}}^0 = \begin{pmatrix} (V_A+1)\mathbb{1}_2 & Z_d\,\sigma_z \\ Z_d\,\sigma_z & (V_A+1)\mathbb{1}_2 \end{pmatrix}, \tag{7}$$

where $\sigma_z = \text{diag}(1,-1)$. Here $V_A = 2\alpha^2$ is Alice's modulation variance in the prepare-and-measure version of the protocol, and Z_d is a function of V_A and the dimension d of the protocol. The two-mode squeezed vacuum, corresponding to a modulation on a sphere whose dimension tends to infinity, has the same form with $Z_\infty := Z_{\text{EPR}} = \sqrt{V_A^2 + 2V_A}$. A comparison of Z_1 (four-state protocol), Z_8 with the maximal value Z_{EPR} is displayed in Fig. 1.

One also knows the covariance matrix Γ_{G} of the state ρ_{G} *after* the quantum channel:

$$\Gamma_{\text{G}} = \begin{pmatrix} (V_A+1)\mathbb{1}_2 & \sqrt{T}\,Z_{\text{EPR}}\,\sigma_z \\ \sqrt{T}\,Z_{\text{EPR}}\,\sigma_z & (1+TV_A+T\xi)\mathbb{1}_2 \end{pmatrix}, \tag{8}$$

where T and ξ refer, respectively, to the transmission and excess noise of the channel and can be estimated experimentally [21].

The last part of the argument consists in proving that the covariance matrix Γ_{key} has the same form (or at least can be safely considered to have the same form) as Γ_{G} after the quantum channel, if one simply replaces Z_{EPR} by Z_d:

$$\Gamma_{\text{key}} = \begin{pmatrix} (V_A+1)\mathbb{1}_2 & \sqrt{T}\,Z_d\,\sigma_z \\ \sqrt{T}\,Z_d\,\sigma_z & (1+TV_A+T\xi)\mathbb{1}_2 \end{pmatrix}. \tag{9}$$

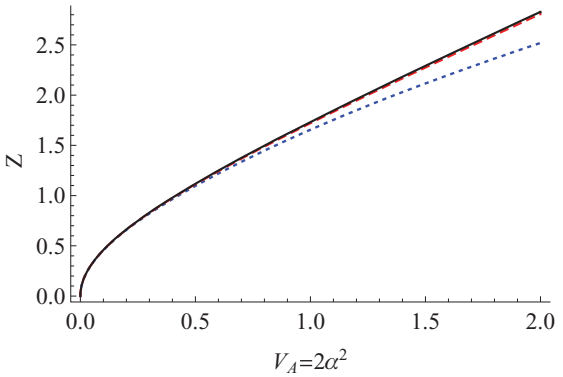

FIG. 1. (Color online) Comparison of the covariance coefficient Z for the states ρ^0_{EPR} (top solid black curve) and ρ^0_{key} for the four-state protocol (bottom short-dash blue curve) and the eight-dimensional protocol (middle long-dash red curve) as a function of the variance of modulation V_A.

This is clear if the quantum channel $\mathcal{T}$ is linear, for instance, Gaussian; but the argument is more involved in the case of an arbitrary quantum channel. In fact, if the channel is linear, Alice and Bob can directly compute the covariance matrix Γ_{key} from the data corresponding to the modulation $p_{\mathrm{key}}(\alpha)$, and the Gaussian modulation or the decoy states are not required in that case [18,19].

Because the modulation considered here is indistinguishable from a Gaussian modulation, the parameter estimation is performed in such a way that there are no privileged direction in phase space. However, this is formally not enough to warrant that Eve's attack has the same symmetry. We therefore provide the following two strategies: either one assumes that the symmetry is not broken by Eve (which in theory could be checked by performing a tomography of the state), or one does not want to make such an assumption and prefers to actively symmetrize the protocol.

In the first scenario, under the assumption that the symmetry of the quantum channel is not broken, the security protocol with decoy states presented here is at least as good as in the case where the channel is linear. Indeed the quantum state shared by Alice and Bob is invariant under the group of conjugate passive symplectic operations applied on Alice's n modes and Bob's n modes, which means that their state can be safely considered to be Gaussian if the analysis is restricted to collective attacks [6].

However, if one does not want to rely on the assumption that the symmetry is not broken, it is possible to remove this assumption thanks to an active symmetrization of the protocol. This is described in detail in Appendixes F and G. This additional step shows that the state ρ_{G} after the quantum channel is rotationally invariant in phase space. Therefore, when restricting ourselves to collective attacks, we conclude, using the technique presented in Ref. [6], that the state ρ_{G} can be safely considered to be Gaussian. In particular, the covariance matrix given in Eq. (9) can be used for the security analysis, with the same values of T and ξ as the ones in Γ_{G}, obtained from the parameter estimation step.

Finally, using the covariance matrix Γ_{key}, one can compute the quantity $\chi(B;E)$ using, for instance, the formalism detailed in [8].

VI. PERFORMANCE OF THE PROTOCOLS

The (asymptotic) secret key rate of the protocols reads

$$K = \beta I(A;B) - \chi(B;E). \tag{10}$$

The idea in our protocols is to use a modulation scheme which is compatible with a very efficient reconciliation, thereby greatly increasing the quantity $\beta I(A;B)$ in comparison with Gaussian modulation protocols.

The price to pay is that this non-Gaussianity makes our bound on $\chi(B;E)$ less tight. This is because the correlation Z_d of the state ρ^0_{key} is strictly less than Z_{EPR} for a given variance of modulation. Interestingly, this discrepancy can be interpreted in terms of excess noise: the fact that ρ^0_{key} displays smaller correlations than the two-mode squeezed state has the same effect as some virtual excess noise. In particular, the value of $\chi(B;E)$ one obtains in the d-dimensional protocol corresponds to the value one would obtain for a Gaussian modulation (GG02) protocol with a quantum channel characterized by a transmission $T_{\mathrm{G}} = T/F \approx T$, and an excess noise $\xi_{\mathrm{G}} = F\xi + (F-1)V_A \approx \xi + (F-1)V_A$, where $F \equiv (Z_{\mathrm{EPR}}/Z_d)^2$. Since one has $F \approx 1$ for reasonable values of V_A (see Fig. 1), the main effect of the non-Gaussian modulation is the *equivalent excess noise* $\Delta\xi = (F-1)V_A$. Figure 2 displays this equivalent excess noise in the case of the four-state protocol ($d = 1$) and the eight-dimensional protocol.

In state-of-the-art implementations [8,9], the excess noise is typically less than a few percent of the shot noise. This gives an approximate limit for the value of the equivalent excess noise that is acceptable. In particular, for the four-state protocol, one needs to work with modulation variances below 0.5 units of shot noise. On the contrary, it becomes possible to work with much higher variances in the case of eight-dimensional protocol.

In Fig. 3, we display the asymptotic secret key rate of the four-state and the eight-dimensional protocols. The various

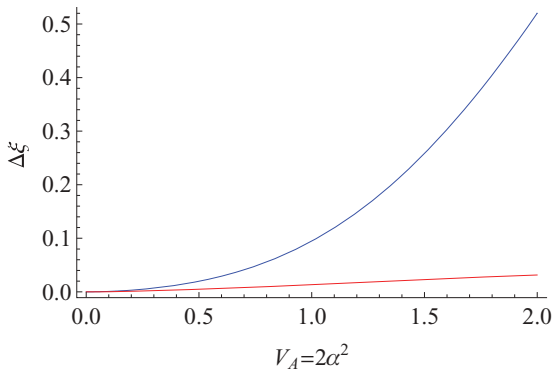

FIG. 2. (Color online) Equivalent excess noise $\Delta\xi$ due to the non-Gaussian modulation as a function of the variance of modulation V_A. Upper curve refers to the four-state protocol, lower curve to the eight-dimensional protocol. By definition, a protocol with a Gaussian modulation does not display any equivalent excess noise. An excess noise of one unit of shot noise corresponds to an entanglement-breaking channel, therefore no security is possible with such a level of noise. This figure clearly shows that the eight-dimensional protocol outperforms significantly the four-state protocol.

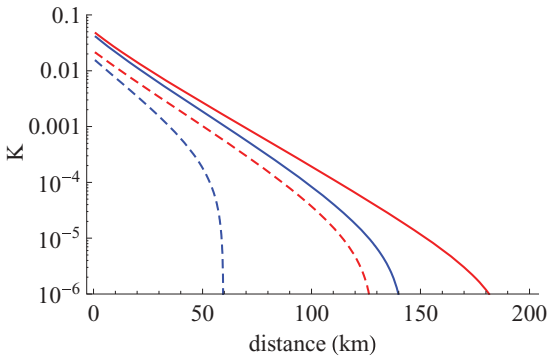

FIG. 3. (Color online) Asymptotic secret key rate K for the eight-dimensional protocol (solid lines) and the four-state protocol (dashed lines) as a function of the distance (assuming transmission through a standard telecommunications fiber with 0.2 dB of loss per kilometer). The various parameters are an excess noise of 0.005 [upper (red) lines] or 0.01 [lower (blue) lines] and a quantum efficiency of the detectors η of 60%. Reconciliation efficiency is supposed to be a conservative 80%.

parameters are chosen conservatively: a quantum efficiency of 60%, a reconciliation efficiency of 80%, and an excess noise of 0.005 or 0.01 units of shot noise. The superiority of the eight-dimensional protocol is quite clear: the secret key rate is higher by nearly an order of magnitude, and one can work with significantly larger modulation variances (the optimized variances are $V_A = 0.3$ for the four-state protocol and $V_A = 0.7$ for the eight-dimensional protocol).

To confirm the robustness of the eight-dimensional protocol, we display in Fig. 4 the secret key rate when finite-size effects are taken into account. The secret key rate is computed against collective attacks, as detailed in Ref. [13]. Among various finite-size effects [22], the most crucial ones for continuous-variable protocols are clearly the imperfect reconciliation efficiency (which prevents the protocol with a Gaussian modulation to achieve key distribution over large distances) and the parameter estimation. While the reconciliation efficiency is taken care of by the eight-dimensional continuous modulation, the parameter estimation is quite sensitive for continuous-variable protocols. In fact, the real problem lies in the estimation of the excess noise ξ, which is very small compared to the shot noise, and thus hard to evaluate accurately.

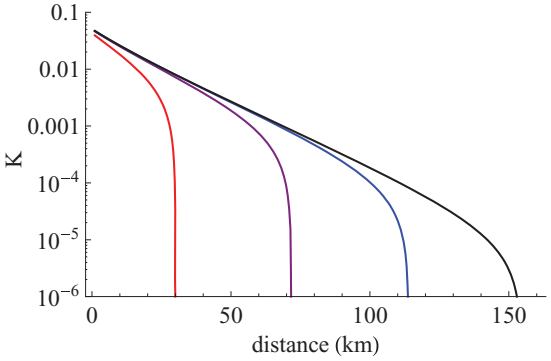

FIG. 4. (Color online) Nonasymptotic secret key rate K for the eight-dimensional protocol, obtained for realistic values: excess noise $\xi = 0.005$, security parameter $\epsilon \approx 10^{-10}$, quantum efficiency of the detectors $\eta = 60\%$, reconciliation efficiency 80% for the BI-AWGN channel, and transmission through a telecommmunications fiber with 0.2 dB of loss per kilometer. Half the samples are used for parameter estimation. From left to right, the block length is equal to 10^8, 10^{10}, 10^{12}, and 10^{14}.

In Fig. 4, all such finite-size effects are taken into account [13]. The results are rather pessimistic, but remember that this is also true for all discrete-variable protocols [23], and the protocols presented here perform reasonably well in comparison. While exchanging 10^{14} quantum signals is rather unrealistic, exchanging 10^{10} or even 10^{11} signals is not completely out of reach of today's technology. Hence, our protocols allow the distribution of secret keys over distances of the order of 100 km, taking into account all finite-size effects.

VII. CONCLUSION AND PERSPECTIVES

In conclusion, we introduced continuous-variable QKD protocols with non-Gaussian modulations. We established the security of these protocols against arbitrary collective attacks, which implies their unconditional security in the asymptotic limit.

The four-state (eight-dimensional) protocol appears optimal *from a practical point of view* among all protocols using a homodyne (heterodyne) detection in the sense that it allows an efficient reconciliation while remaining as close as possible to the theoretically optimal Gaussian protocols [3,14].

The main open questions concern the status of the decoy states. First, it should be possible to prove the security of the protocols considered here without requiring any decoy states. This might come at the price of slightly degraded bounds (to take into account possible non-Gaussian attacks). Second, without removing the decoy states, an important question is how well they should approximate the Gaussian distribution. Put otherwise, how indistinguishable should the distribution corresponding to key distillation be from the one used for parameter estimation? In particular, how should this distinguishability be taken into account in the overall security parameter of the protocol?

Finally, the most outstanding problem that remains for continuous-variable QKD protocol concerns general attacks. This question was partially answered with the derivation of a de Finetti–type theorem for quantum systems of infinite dimension [7]. However, the bounds obtained there are not good enough to be used in practice. Hence, it seems crucial to see if the postselection technique introduced in [24] can be adapted to continuous variables, since this technique is already known to provide much better (almost tight) bounds than the de Finetti theorem in the case of discrete variables [25].

ACKNOWLEDGMENTS

The authors acknowledge fruitful discussions with Frédéric Grosshans, Norbert Lütkenhaus, and Renato Renner. This work was carried out in the framework of the ANR project SEQURE (ANR-07-SESU-011-01). A.L. received financial support from the EU ERC Starting Grant PERCENT.

APPENDIX A: LINEAR QUANTUM CHANNELS

We shall define a linear quantum channel by the input-output relations of the quadrature operators in Heisenberg representation:

$$\begin{aligned} X_{\text{out}} &= g_X X_{\text{in}} + B_X, \\ P_{\text{out}} &= g_P P_{\text{in}} + B_P, \end{aligned} \tag{A1}$$

where the added noises B_X, B_P are uncorrelated with the input quadratures X_{in}, P_{in}. Such relations have been extensively used, for instance, in the context of quantum nondemolition (QND) measurements of continuous variables [26], and they are closely related to the linearized approximation commonly used in quantum optics. Gaussian channels (channels that preserve the Gaussianity of the states) are usual examples of linear quantum channels. However, linear quantum channels may also be non-Gaussian, this will be the case, for instance, if the added noises B_X, B_P are non-Gaussian.

For our purpose, the main advantage of a linear quantum channel is that it will be characterized by transmission coefficients $T_X = g_X^2$ and $T_P = g_P^2$, and by the variances of the added noises B_X and B_P. These quantities can be determined even if the modulation used by Alice is non-Gaussian, with the same measured values as when the modulation is Gaussian (because these values are intrinsic properties of the channel). The relevant covariance matrix can then be easily determined, and Eve's information can be bounded by using the Gaussian optimality theorem. This justifies the approach taken in Refs. [16,18], but unfortunately this is not fully general, contrary to the proof of the present paper.

APPENDIX B: EFFICIENT REVERSE RECONCILIATION

The goal of this Appendix is to explain how an efficient reconciliation can be achieved for the various modulation schemes considered in this article, for arbitrarily low SNR. Reconciliation in a QKD protocol is very similar to the problem of channel coding (that is, transmitting information efficiently and reliably on a noisy communication channel) with the additional constraint that the input of the channel is not chosen by Alice but instead randomly picked from a given probability distribution corresponding to the modulation scheme. In particular, the usual task is the following: Alice and Bob are given two n-dimensional real vectors $\mathbf{x}$ and $\mathbf{y}$ and their goal is to agree on a common bit string $\mathbf{u}$. A supplementary constraint when dealing with reverse reconciliation is that $\mathbf{u}$ should be a function (possibly randomized) of $\mathbf{y}$, and that all public communication should be from Bob to Alice.

There exists a standard technique for reducing the problem of reconciliation to the one of channel coding, namely, coset coding introduced by Wyner [27]. The idea is that Bob will use an additional public (but authenticated) channel to describe a function f such $f(\mathbf{y}) = \mathbf{u}$. Alice can then apply this function to her vector and obtain $\mathbf{v} := f(\mathbf{x})$. In the case of coset coding, the description of f is simply a translation corresponding to the syndrome of y for a linear error correcting code C. This gives rise to a virtual communication channel with input $\mathbf{u}$ and output $\mathbf{v}$ for which one can apply standard channel coding techniques.

One case where reconciliation can be performed very efficiently occurs when the virtual channel is a binary-input additive white Gaussian noise (BI-AWGN) channel, meaning that the coordinates of $\mathbf{u}$ and $\mathbf{v}$ are related through

$$v_i = u_i + w_i, \tag{B1}$$

where $u_i \in \{-1,1\}$ and w_i is a centered normal random variable.

For experimental realizations of continuous-variable QKD, the quantum channel always behaves in very good approximation like a Gaussian channel, and the BI-AWGN channel is therefore the model of interest here. In this case, it is possible to show that the existence of an efficiently computable function f is possible only for very specific modulation schemes, namely, the cases where d-uplets of x_i are distributed uniformly on the unit sphere in dimension 1, 2, 4, or 8 [10]. The case $d = 1$ corresponds to a binary modulation, that is, the four-state protocol (which indeed displays a binary modulation for *each* quadrature); the case $d = 8$ corresponds to the eight-dimensional modulation.

Let us therefore consider d-uplets $\mathbf{x^d}$ and $\mathbf{y^d}$, with $\mathbf{x^d} \sim \mathcal{U}(\mathcal{S}^{d-1})$ and $\mathbf{y^d} = \mathbf{x^d} + \mathbf{z^d}$ with $\mathbf{z^d} \sim \mathcal{N}(0,\sigma^2)^d$. Without loss of generality, we restrict our attention to the case where the transmission is 1. For dimensions 1, 2, 4, and 8, the unit sphere $\mathcal{S}^{d-1}$ has a division algebra. In particular, a d-dimensional vector can be identified with an element of $\mathbb{R}^d$, that is, a real number ($d = 1$), a complex number ($d = 2$), a quaternion ($d = 4$), or an octonion ($d = 8$). Therefore, both multiplication and division are well defined in this context.

Bob chooses a random element $\mathbf{u^d} \in \{-1/\sqrt{d}, 1/\sqrt{d}\}^d$ with the uniform distribution on the d-dimensional hypercube and sends the variable $\mathbf{t^d} := \mathbf{u^d}\mathbf{y^d}$ to Alice (through the classical channel). Alice computes $\mathbf{v^d} := \mathbf{t^d}\left(\mathbf{x^d}\right)^{-1}$ which is possible because $\mathcal{S}^{d-1}$ is a division algebra. We now wish to prove that the channel $\mathbf{u^d} \to \mathbf{v^d} = \mathbf{u^d} + \mathbf{w^d}$ is a BI-AWGN channel. Let us characterize the noise $\mathbf{w^d}$ on this virtual channel:

$$\begin{aligned} \mathbf{w^d} &\equiv \mathbf{v^d} - \mathbf{u^d}, && \text{(B2)} \\ &= \mathbf{t^d}(\mathbf{x^d})^{-1} - \mathbf{u^d}, && \text{(B3)} \\ &= \mathbf{u^d}\mathbf{y^d}(\mathbf{x^d})^{-1} - \mathbf{u^d}, && \text{(B4)} \\ &= \mathbf{u^d}(\mathbf{y^d}(\mathbf{x^d})^{-1} - 1), && \text{(B5)} \\ &= \mathbf{u^d}[(\mathbf{x^d} + \mathbf{z^d})(\mathbf{x^d})^{-1} - 1], && \text{(B6)} \\ &= \mathbf{u^d}\mathbf{z^d}(\mathbf{x^d})^{-1}. && \text{(B7)} \end{aligned}$$

Since $\mathbf{u^d}$ and $(\mathbf{x^d})^{-1}$ are simply rotations on $\mathcal{S}^{d-1}$, one concludes that $\mathbf{w^d} \sim \mathcal{N}(0,\sigma^2)^{\otimes d}$, which proves that the virtual channel $\mathbf{u^d} \to \mathbf{v^d}$ is indeed a BI-AWGN channel, for which efficient error-correcting codes are available.

If $n = d \times m$, Alice and Bob simply divide their data into m d-uplets and proceed as described above. All that is left to do is to use coset coding to finish the reconciliation: Bob sends the syndrome of $\mathbf{u^n}$ for a linear code Alice and he agreed on beforehand. Alice simply decodes her word $\mathbf{v^n}$ in the coset code defined by the syndrome. This can be done very efficiently with capacity approaching codes such as low-density parity-check (LDPC) codes [28]. If the SNR is very low, then one can work with a concatenation of a low-rate LDPC code (such as a multi-edge LDPC code, for instance [29]) with a repetition code. This

simple technique allows one to obtain good error-correcting codes of arbitrarily low rate (see Chap. 5.2 of Ref. [30] for more details concerning this concatenation technique).

APPENDIX C: DECOYS WITH COHERENT STATES

In the entanglement-based version of the protocol, one has to apply the generalized measurements $\{\Pi_d, \mathbb{1} - \Pi_d\}$. In the prepare-and-measure scenario, it is therefore necessary for Alice to send states which are compatible with these measurements. The states corresponding to the operator Π_d are not a problem, since by construction, they correspond to the modulation used to distill the key, that is, coherent states drawn uniformly on the sphere in d dimensions. The states corresponding to the operator $\mathbb{1} - \Pi_d$ might be a little bit more problematic in the sense that they are not usually easy to produce experimentally. Ideally, one would like to be able to produce these states simply by modulating coherent states (which are the only states simple enough to allow for a realistic QKD protocol). In particular, if Alice applies the measurements $\{\Pi_d, \mathbb{1} - \Pi_d\}$, then one obtains the following relation:

$$\sigma_{\mathrm{G}}^{d} = p\,\sigma_{\mathrm{key}}^{d} + (1-p)\,\sigma_{\mathrm{decoy}}^{d}, \tag{C1}$$

with $p = p_d^{\mathrm{succ}}$. However, in this case, the state $\sigma_{\mathrm{decoy}}^{d}$ does not have a positive P function, meaning that it cannot be obtained as a mixture of coherent states.

We now present two different techniques to deal with this problem: either one replaces $\sigma_{\mathrm{key}}^{d}$ with a noisy version (see Appendix C1) or one relaxes Eq. (C1) and considers instead an approximate version of the decoy states (see Appendix C2).

1. *Perfect decoys with noisy signal*

Let us consider the prepare-and-measure version of the d-dimensional protocol. In this case, the Gaussian modulation can be seen as sending $d/2$ coherent states $|\alpha_1 + i\alpha_2\rangle, \ldots, |\alpha_{d-1} + i\alpha_d\rangle$ such that the random variables α_i are independent and identically distributed centered normal variables. Without loss of generality, we consider variables with variance $1/d$. Taking advantage of the rotational invariance of the Gaussian distribution, one can equivalently choose the random vector $\alpha := (\alpha_1, \ldots, \alpha_d)$ by first picking uniformly a random point of the unit sphere $\mathcal{S}^{d-1}$ in d dimensions, and drawing the radius $r := \sqrt{\sum \alpha_i^2}$ of the vector α from a chi distribution with d degrees of freedom. In particular, the probability density function of r is

$$f(r,d) = \frac{2(d/2)^{d/2} r^{d-1} e^{-dr^2/2}}{(d/2-1)}. \tag{C2}$$

The probability distributions corresponding to $d = 1$, 2, 4, and 8 are displayed in Fig. 5. In particular, it should be noted that they become more and more peaked as the dimension d increases. In this picture, the probability distribution corresponding to the key, that is, $\sigma_{\mathrm{key}}^{d}$, is by construction a Dirac distribution centered in 1.

The first approach we investigate aims at satisfying Eq. (C1) exactly while allowing for a positive (and non-negligible) probability p of sending a signal state. This is done by considering slightly noisy versions of the true signal modulation.

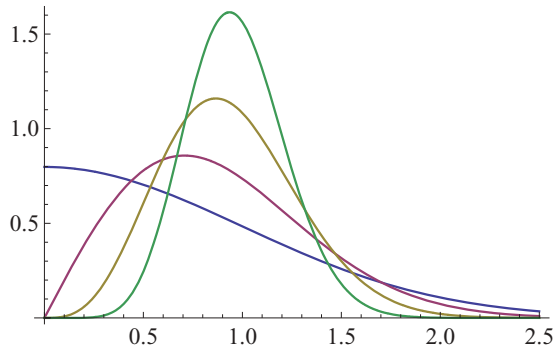

FIG. 5. (Color online) Probability density functions for the radius of α in the d-dimensional protocol for a Gaussian modulation. From least peaked to most peaked, $d = 1$, 2, 4, and 8.

In particular, one chooses two parameters $\gamma_{\min} \in [0,1]$ and $\gamma_{\max} \geqslant 1$ and defines the states used for key distillation as the ones with a radius bounded by these two parameters: $\gamma_{\min} \leqslant r \leqslant \gamma_{\max}$. The decoy states then simply correspond to the remaining states. Provided that $\gamma_{\min}$ and $\gamma_{\max}$ are close enough to 1, the penalty imposed by this noise, compared to the ideal case where the key modulation is strictly equal to 1, is negligible in terms of reconciliation efficiency.

On the other hand, one should not choose values too close to 1, otherwise the probability p that a given state can be used for key distillation will be very small:

$$p = \int_{\gamma_{\min}}^{\gamma_{\max}} f(r,d)\,dr. \tag{C3}$$

Hence, optimizing the values of $\gamma_{\min}$ and $\gamma_{\max}$ should be seen as a tradeoff between the penalty imposed on reconciliation efficiency and the probability that a given state can be used for key distillation.

Note also that with this approach, the state $\sigma_{\mathrm{key}}^{d}$ is not longer described as in Appendix E. In particular, the covariance matrix of the new state is a little bit different from the one presented in Appendix E. We do not give an explicit derivation of the new covariance matrix here, but we point out that because the new state used for the key is actually closer to a Gaussian modulation, the Holevo information between Bob and Eve can still be safely bounded using the covariance matrix given in Eq. (9).

We now give a second approach to the problem of approximating decoy states with coherent states.

2. *Approximate decoys with noiseless signal*

Our goal is still to achieve the following equality:

$$\sigma_{\mathrm{G}}^{d} = p\,\sigma_{\mathrm{key}}^{d} + (1-p)\,\sigma_{\mathrm{decoy}}^{d}, \tag{C4}$$

but this time, without considering a noisy version of $\sigma_{\mathrm{key}}^{d}$. If one chooses $p = p_d^{\mathrm{succ}}$ as defined in Appendix D, then the state $\sigma_{\mathrm{decoy}}^{d}$ does not have a positive P function, meaning that it cannot be obtained as a mixture of coherent states. Fortunately, P functions can be regularized rather well, and for our purpose, it is sufficient to find a state $\sigma_{\mathrm{decoy}}^{d}$ with a non-negative P function such that Eq. (C4) only holds *approximately*. Here,

approximately should be understood in terms of the trace distance.

More precisely, we are interested in finding a probability distribution (hence non-negative) $P(\alpha)$ such that

$$\left\| \sigma_{\mathrm{G}}^{d} - p\,\sigma_{\mathrm{key}}^{d} - (1-p)\int P(\alpha)|\alpha\rangle\langle\alpha|\,\mathrm{d}\alpha \right\|_1 \leqslant \epsilon \qquad \text{(C5)}$$

for a value of ϵ sufficiently small. If Eq. (C5) holds and Alice uses the modulation $P(\alpha)$ for the decoy states, then the probability that Eve can distinguish the states used for key distillation from the ones used for parameter estimation is upper bounded by ϵ.

Here, we do not give a solution for the problem of finding the best distribution P compatible with a success probability p and the error ϵ, but we point out that the usual optimization tools (for instance, the optimization toolbox of MATLAB) allow one to find very good instances of P. For example, for the two-dimensional protocol and $\alpha = 0.5$, we could obtain a value of ϵ less than 10^{-5} for $p = 1/2$ with a distribution P corresponding to a sum of six Dirac distributions, that is, a mixture of six particular coherent states.

The natural question that arises here is, how good should the approximation be? Is a trace distance equal to 10^{-5} sufficient to guarantee a reasonable level of security? Or should one aim for a value of 10^{-10}?

Note that if the approximation is not perfect, it means that Eve might have a (very) small probability to distinguish between the states used for the key distillation and those used for parameter estimation. However, discriminating between these two modulations does not appear to be a good solution for Eve, as this would induce a lot of phase noise in the signal. Indeed, because all the modulations considered here are phase invariant, the optimal discrimination procedure consists in projecting the states onto Fock states, thereby erasing all the phase information. For this reason, a trace distance of 10^{-5} between ρ_{decoy} and its approximation is very likely to be sufficient for any practical implementation.

APPENDIX D: MEASUREMENT OPERATOR

We now describe the general measurement $\{\Pi_d, \mathbb{1} - \Pi_d\}$ performed by Alice to prepare the state ρ_{key}^{0} from a two-mode squeezed vacuum. The state to which this measurement is applied is a (possibly multimodal) two-mode squeezed vacuum as described in Table I.

TABLE I. Parametrization of the various protocols. Parameter d corresponds to the number of quadratures that should be processed together.

d	Protocol	Resource state
1	4-state	$\|\mathrm{EPR}\rangle$
2	2-dim	$\|\mathrm{EPR}\rangle$
4	4-dim	$\|\mathrm{EPR}\rangle^{\otimes 2}$
8	eight-dim	$\|\mathrm{EPR}\rangle^{\otimes 4}$

1. Four-state protocol: $d = 1$

The operator Π_1 is defined as $\Pi_1 = M_1^{\dagger} M_1$ with

$$M_1 = m_1 \sum_{k=0}^{3} |\psi_k\rangle\langle e_k|, \qquad \text{(D1)}$$

and

$$m_1 := \frac{e^{(1+\alpha^2)/2}}{2}\sqrt{\frac{\lfloor 1+\alpha^2 \rfloor !}{(1+\alpha^2)^{\lfloor 1+\alpha^2 \rfloor}}}. \qquad \text{(D2)}$$

The states $|\psi_k\rangle$ are defined in Appendix E, and

$$|e_k\rangle = e^{-\beta^2/2}\sum_{n=0}^{\infty}\frac{\beta_k^{*n}}{\sqrt{n!}}|n\rangle, \qquad \text{(D3)}$$

with $\beta = \sqrt{1+\alpha^2}$ and $\mathrm{Arg}(\beta_k) = \mathrm{Arg}(\alpha_k)$.

When performing the general measurement $\{\Pi_1, \mathbb{1} - \Pi_1\}$, conditioned on the result corresponding to Π_1, the state ρ is transformed into ρ':

$$\rho \longrightarrow \rho' := \frac{M_1 \rho M_1^{\dagger}}{\mathrm{tr} M_1 \rho M_1^{\dagger}}. \qquad \text{(D4)}$$

Let us consider the state $\rho_{\mathrm{G}}^{0} := |\mathrm{EPR}\rangle\langle\mathrm{EPR}|$ with

$$|\mathrm{EPR}\rangle = \sum_{n=0}^{\infty}\sqrt{\frac{\alpha^{2n}}{(1+\alpha^2)^{n+1}}}|n\rangle|n\rangle. \qquad \text{(D5)}$$

Conditioned on the result Π_1, one obtains

$$M_1 \rho_{\mathrm{G}} M_1^{\dagger} = \frac{4m_1^2 e^{-(\alpha^2+1)}}{\alpha^2+1}|\Psi_1\rangle\langle\Psi_1|, \qquad \text{(D6)}$$

with

$$|\Psi_1\rangle = \frac{1}{2}\sum_{k}^{3}|\psi_k\rangle|\alpha_k\rangle. \qquad \text{(D7)}$$

The condition $\Pi \leqslant \mathbb{1}$ leads to $m_1 \leqslant m_1^{\max}$ with

$$m_1^{\max} := \frac{e^{(1+\alpha^2)/2}}{2}\sqrt{\frac{\lfloor 1+\alpha^2 \rfloor !}{(1+\alpha^2)^{\lfloor 1+\alpha^2 \rfloor}}}. \qquad \text{(D8)}$$

The probability of obtaining the result corresponding to Π_1, meaning successfully creating a state $|\Psi_1\rangle$, is

$$p_1^{\mathrm{succ}} = \mathrm{tr} M_1 \rho M_1^{\dagger} = \frac{\lfloor 1+\alpha^2 \rfloor !}{(1+\alpha^2)^{\lfloor 2+\alpha^2 \rfloor}} \qquad \text{(D9)}$$

and is displayed in Fig. 6 as a function of α.

2. Continuous modulations: $d = 2,4,8$

For $d = 2,4,8$, one has

$$|\Psi_d\rangle = \sum_{k=0}^{\infty}\sqrt{f_n(k)}\,|\psi_k^d\rangle, \qquad \text{(D10)}$$

where

$$f_n(k) = e^{-n\alpha^2}\frac{n^k \alpha^{2k}}{k!}, \qquad \text{(D11)}$$

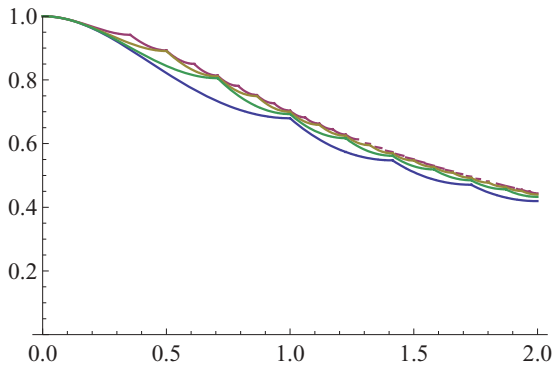

FIG. 6. (Color online) Probability of success of the measurement Π_d as a function of α. From bottom to top, $d = 1$, 2, 4, and 8.

and

$$\left|\psi_k^2\right\rangle := |k\rangle|k\rangle, \quad \left|\psi_k^4\right\rangle := \frac{1}{\sqrt{k+1}} \sum_{k_1=0}^{k} |k_1, k-k_1\rangle|k_1, k-k_1\rangle,$$

$$\left|\psi_k^8\right\rangle := \frac{1}{\sqrt{\binom{k+3}{3}}} \sum_{\sum_i k_i = k} |k_1,k_2,k_3,k_4\rangle|k_1,k_2,k_3,k_4\rangle.$$

One also has

$$|\mathrm{EPR}\rangle = \sum_{k=0}^{\infty} \sqrt{g_2(k)}\left|\psi_k^2\right\rangle, \quad |\mathrm{EPR}\rangle^{\otimes 2} = \sum_{k=0}^{\infty} \sqrt{g_4(k)}\left|\psi_k^4\right\rangle,$$

$$|\mathrm{EPR}\rangle^{\otimes 4} = \sum_{k=0}^{\infty} \sqrt{g_8(k)}\left|\psi_k^8\right\rangle,$$

with

$$g_n(k) = \binom{n+k-1}{k} \frac{\alpha^{2k}}{(1+\alpha^2)^{n+k}}. \tag{D12}$$

g_n is a negative binomial distribution $B_N(d, \frac{\alpha^2}{1+\alpha^2})$. Let us define the operators Π_2, Π_4, and Π_8 as

$$\Pi_d = \pi_d \sum_{k=0}^{\infty} \frac{f_d(k)}{g_d(k)} \mathrm{tr}_B\left|\psi_k^d\right\rangle\left\langle\psi_k^d\right|, \tag{D13}$$

where π_d is given by

$$\pi_d(\alpha) = \min_{k\in\mathbb{N}} \frac{g(k)}{f(k)}, \tag{D14}$$

ensuring that Π_d is a genuine positive operator-valued measure (POVM) element. It is straightforward to check that the probability of success of the measurement is

$$p_d^{\mathrm{succ}} = \frac{g(\lceil \alpha^2 d \rceil)}{f(\lceil \alpha^2 d \rceil)}. \tag{D15}$$

APPENDIX E: COVARIANCE MATRICES

In this Appendix, we derive the covariance matrices of the states corresponding to the four-state protocol, which is optimal with a homodyne detection, and the eight-dimensional protocol, which is optimal for a heterodyne detection. The covariance matrices corresponding to the other (suboptimal) choices of modulation can be found with a similar technique.

Let us note $|\Psi_d\rangle$ for $d \in \{1,2,4,8\}$ the state used for the key distillation in each protocol, i.e., $\rho_{\mathrm{key}}^0 = |\Psi_d\rangle\langle\Psi_d|$; and $|\Psi_1\rangle$ is the initial bipartite state for the four-state protocol, whereas $|\Psi_8\rangle$ corresponds to the eight-dimensional protocol. The two-mode squeezed vacuum is noted as $|\mathrm{EPR}\rangle$.

1. *Four-state protocol: d = 1*

Let us use the notation $|\alpha_k\rangle := |\alpha e^{(2k+1)i\pi/4}\rangle$ for $k \in \{0,1,2,3\}$ and $\alpha > 0$. In the four-state protocol, the state ρ_{key}^0 is a pure state $|\Psi_1\rangle$ defined as

$$|\Psi_1\rangle = \sum_{k=0}^{3} \sqrt{\lambda_k}|\phi_k\rangle|\phi_k\rangle, \tag{E1}$$

$$= \frac{1}{2}\sum_{k=0}^{3} |\psi_k\rangle|\alpha_k\rangle, \tag{E2}$$

where

$$|\phi_k\rangle = \frac{e^{-\alpha^2/2}}{\sqrt{\lambda_k}} \sum_{n=0}^{\infty} (-1)^n \frac{\alpha^{4n+k}}{\sqrt{(4n+k)!}}|4n+k\rangle, \tag{E3}$$

$$|\psi_k\rangle = \frac{1}{2}\sum_{m=0}^{3} e^{i(1+2k)m\pi/4}|\phi_m\rangle \tag{E4}$$

for $k \in \{0,1,2,3\}$ and

$$\begin{aligned} \lambda_{0,2} &= \tfrac{1}{2}e^{-\alpha^2}[\cosh(\alpha^2) \pm \cos(\alpha^2)], \\ \lambda_{1,3} &= \tfrac{1}{2}e^{-\alpha^2}[\sinh(\alpha^2) \pm \sin(\alpha^2)]. \end{aligned} \tag{E5}$$

Let us note a and b as the annihilation operators on the two modes. Applying a to $|\phi_k\rangle$ gives

$$a|\phi_k\rangle = \alpha \frac{\sqrt{\lambda_{k-1}}}{\sqrt{\lambda_k}}|\phi_{k-1}\rangle \tag{E6}$$

for $k \in \{1,2,3\}$ and

$$a|\phi_0\rangle = -\alpha \frac{\sqrt{\lambda_3}}{\sqrt{\lambda_0}}|\phi_3\rangle. \tag{E7}$$

Let us compute the covariance matrix Γ_1 of the bipartite state $|\Psi_1\rangle$. It has the following form:

$$\Gamma_1 = \begin{pmatrix} X\mathbb{1}_2 & Z_1\sigma_z \\ Z_1\sigma_z & Y\mathbb{1}_2 \end{pmatrix}. \tag{E8}$$

One has

$$X = Y = \langle\Psi_1|1 + 2a^\dagger a|\Psi_1\rangle = \langle\Psi_1|1 + 2b^\dagger b|\Psi_1\rangle, \tag{E9}$$

$$= \mathrm{tr}\left(1 + 2\sum_{k=0}^{3} a^\dagger a\ \lambda_k|\phi_k\rangle\langle\phi_k|\right), \tag{E10}$$

$$= 1 + 2\sum_{k=0}^{3} \lambda_k\langle\phi_k|a^\dagger a|\phi_k\rangle, \tag{E11}$$

$$= 1 + 2\alpha^2 \sum_{k=0}^{3} \lambda_k \frac{\lambda_{k-1}}{\lambda_k}, \tag{E12}$$

$$= 1 + 2\alpha^2. \tag{E13}$$

The correlation term Z_1 of the covariance matrix is given by

$$Z_1 = \langle\Psi_1|ab + a^\dagger b^\dagger|\Psi_1\rangle, \tag{E14}$$
$$= 2\mathrm{Re}\langle\Psi_1|ab|\Psi_1\rangle. \tag{E15}$$

One has

$$ab|\Psi_1\rangle = ab\sum_{k=0}^{3}\sqrt{\lambda_k}|\phi_k\rangle|\phi_k\rangle, \tag{E16}$$
$$= \alpha^2\sum_{k=0}^{3}\frac{\lambda_{k-1}}{\lambda_k}\sqrt{\lambda_k}|\phi_{k-1}\rangle|\phi_{k-1}\rangle, \tag{E17}$$

where addition should be understood modulo 4. Finally, we obtain

$$Z_1 = 2\alpha^2\sum_{k=0}^{3}\frac{\lambda_{k-1}^{3/2}}{\lambda_k^{1/2}}. \tag{E18}$$

It may be noticed that at the lowest order in α, the states $|\phi_k\rangle$ are simply the number states $|k\rangle$ for $k = 0, 1, 2, 3$, which are independent of α. The states $|\psi_k\rangle$ are four orthogonal linear combinations of these four number states, with coefficients of the form $e^{ip\pi/4}$, where p is an integer. In addition, the state $|\Psi_1\rangle$ is simply $(1-\alpha^2/2)|00\rangle + \alpha^2|11\rangle$, which is also the lowest-order (Gaussian) EPR state. Correspondingly, $Z_1 = Z_{\mathrm{EPR}} = 2\alpha = \sqrt{2V_A}$, in the limit where α tends to 0.

Since the entangled state is already Gaussian in this regime, no decoy states are needed, and Ref. [16] can be used directly to establish the unconditional security of the protocol. Unfortunately, this approach is restricted to values of α which are too small to be useful in practice; this is why the more powerful proof presented in the present paper is needed.

2. *Eight-dimensional protocol: d = 8*

The partial trace $\sigma^8_{\mathrm{key}} = \mathrm{tr}_A(|\Psi_8\rangle\langle\Psi_8|)$ is defined by the modulation scheme: it is the uniform mixture of quadrimodal coherent states over a real seven-dimensional sphere:

$$\mathrm{tr}_A(|\Psi_8\rangle\langle\Psi_8|) := \int_{\mathcal{S}_\alpha}|\alpha_1,\alpha_2,\alpha_3,\alpha_4\rangle\langle\alpha_1,\alpha_2,\alpha_3,\alpha_4|dS, \tag{E19}$$

where $|\alpha_1,\alpha_2,\alpha_3,\alpha_4\rangle := |\alpha_1\rangle|\alpha_2\rangle|\alpha_3\rangle|\alpha_4\rangle$ and the sphere $\mathcal{S}_\alpha$ is defined as

$$\mathcal{S}_\alpha \equiv \{(\alpha_1,\alpha_2,\alpha_3,\alpha_4) \in \mathbb{C}^4 \mathrm{s.t.} |\alpha_{4k}|^2 + |\alpha_{4k+1}|^2 + |\alpha_{4k+2}|^2 + |\alpha_{4k+3}|^2 = 4\alpha^2\}, \tag{E20}$$

and dS is the Haar measure on $\mathcal{S}_\alpha$. Because this state is a four-mode orthogonally invariant state (by construction), it can be written as [31]

$$\mathrm{tr}_A(|\Psi_8\rangle\langle\Psi_8|) = \sum_{k=0}^{\infty}\lambda_k\,\sigma_k^4, \tag{E21}$$

where

$$\sigma_k^4 = \frac{1}{\binom{k+3}{3}}\sum_{\substack{k_1\cdots k_4\\ \text{s.t. }\sum_i k_i = k}}|k_1,k_2,k_3,k_4\rangle\langle k_1,k_2,k_3,k_4|. \tag{E22}$$

To determine the $\{\lambda_k\}_{k=0,\ldots,\infty}$, we compute the probability $\Pr(k)$ of finding k photons in the four-mode state $\mathrm{tr}_A(|\Psi_8\rangle\langle\Psi_8|)$:

$$\Pr(k) = \mathrm{tr}\left[\mathrm{tr}_A(|\Psi_8\rangle\langle\Psi_8|)\sigma_k^4\right], \tag{E23}$$
$$= \langle 2\alpha|\langle 0|\langle 0|\langle 0|\,\sigma_k^4\,|2\alpha\rangle|0\rangle|0\rangle|0\rangle, \tag{E24}$$

since $|2\alpha\rangle|0\rangle|0\rangle|0\rangle \in \mathcal{S}_4$. Because the coherent state $|0\rangle$, which refers to the vacuum, does not contain a photon, one has

$$\Pr(k) = \langle 2\alpha|\sigma_k^4|2\alpha\rangle, \tag{E25}$$
$$= e^{-4\alpha^2}\frac{(2\alpha)^{2k}}{k!}, \tag{E26}$$
$$= \lambda_k. \tag{E27}$$

We therefore get the expression

$$\mathrm{tr}_A(|\Psi_8\rangle\langle\Psi_8|) = e^{-4\alpha^2}\sum_{k=0}^{\infty}\frac{(2\alpha)^{2k}}{k!}\,\sigma_k^4. \tag{E28}$$

Finally $|\Psi_8\rangle$ is defined as

$$|\Psi_8\rangle := e^{-2\alpha^2}\sum_{k=0}^{\infty}\frac{(2\alpha)^k}{\sqrt{k!}}\,|\psi_k^4\rangle, \tag{E29}$$

where

$$|\psi_k^4\rangle = \frac{1}{\sqrt{\binom{k+3}{3}}}\sum_{\substack{k_1\cdots k_4\\ \text{s.t. }\sum_i k_i = k}}|k_1,k_2,k_3,k_4\rangle|k_1,k_2,k_3,k_4\rangle. \tag{E30}$$

Let us compute the covariance matrix Γ_8 of $|\Psi_8\rangle$. It has the form

$$\Gamma_8 = \bigoplus_{i=1}^{4}\begin{pmatrix} X\mathbb{1}_2 & Z_8\sigma_z \\ Z_8\sigma_z & X\mathbb{1}_2\end{pmatrix}, \tag{E31}$$

where

$$X = \langle\Psi_8|1 + 2a_1^\dagger a_1|\Psi_8\rangle = \langle\Psi_8|1 + 2b_1^\dagger b_1|\Psi_8\rangle, \tag{E32}$$
$$Z_8 = \langle\Psi_8|a_1b_1 + a_1^\dagger b_1^\dagger|\Psi_8\rangle, \tag{E33}$$

where a_1 and b_1 refer to Alice and Bob's annihilation operators relative to the first mode.

Tracing $|\psi_k^4\rangle$ over the last three modes gives ρ_k^1

$$\rho_k^1 = \frac{1}{\binom{k+3}{3}}\sum_{l=0}^{k}\binom{k-l+2}{2}|l,l\rangle\langle l,l|. \tag{E34}$$

One immediately has

$$\langle\Psi_8|a_1^\dagger a_1|\Psi_8\rangle = \frac{1}{\binom{k+3}{3}}\sum_{l=0}^{k}l\binom{k-l+2}{2} = \frac{k}{4}. \tag{E35}$$

Then,

$$\mathrm{tr}(a_1^\dagger a_1\rho^4) = \sum_{k=0}^{\infty}e^{-4\alpha^2}\frac{(2\alpha)^{2k}}{k!}\frac{k}{4}, \tag{E36}$$
$$= \alpha^2, \tag{E37}$$

which gives $X = 1 + 2\alpha^2$.

Let us now compute $Z_8 = \langle\Psi_8|a_1b_1 + a_1^\dagger b_1^\dagger|\Psi_8\rangle$. First, one notes that $\langle\phi_l^4|a_1b_1|\psi_k^4\rangle = 0$ except if $l = k-1$. Some

combinatorics show that

$$\langle \phi_{k-1}^4 | a_1 b_1 | \psi_k^4 \rangle = \frac{1}{\sqrt{\binom{k+3}{3}\binom{k+2}{3}}} \sum_{l=0}^{k} l \binom{k-l+2}{2}, \quad \text{(E38)}$$

$$= \frac{1}{4}\sqrt{k(k+3)}. \quad \text{(E39)}$$

Using the expression of $|\Psi_8\rangle$, one obtains

$$\langle \Psi_8 | a_1 b_1 | \Psi_8 \rangle = \frac{1}{4} e^{-4\alpha^2} \sum_{k=0}^{\infty} \frac{\sqrt{k+4}}{k!} (2\alpha)^{2k+1}, \quad \text{(E40)}$$

and finally

$$Z_8 = \frac{1}{2} e^{-4\alpha^2} \sum_{k=0}^{\infty} \frac{\sqrt{k+4}}{k!} (2\alpha)^{2k+1}. \quad \text{(E41)}$$

APPENDIX F: SYMMETRIZATION OF THE PROTOCOL

A quantum-key-distribution protocol is described as a map $\mathcal{E}$ [24]:

$$\mathcal{E} : \rho_{AB} \longmapsto (\mathcal{S}_A, \mathcal{S}_B, \mathcal{C}), \quad \text{(F1)}$$

where $\rho_{AB} \in (\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n}$ is the n-mode bipartite state shared by Alice and Bob at the end of the distribution phase (in the entanglement-based protocol), S_A and S_B are, respectively, Alice and Bob's final keys, and $\mathcal{C}$ is a transcript of all classical communication as well as Alice and Bob's raw data.

A protocol $\mathcal{E}$ is said to be invariant under some set of transformations $\mathcal{G}$ if for any element $g \in \mathcal{G}$, there exists a completely positive trace-preserving (CPTP) map $\mathcal{K}_g$ such that

$$\mathcal{E} \circ g = \mathcal{K}_g \circ \mathcal{E}. \quad \text{(F2)}$$

Let us consider the uniform measure $\mu_{\mathcal{G}}$ on $\mathcal{G}$. If the protocol $\mathcal{E}$ is invariant under the set $\mathcal{G}$, then it is sufficient to prove the security of $\mathcal{E}$ for states displaying the same symmetry. In particular, it is sufficient to consider states of the form

$$\bar{\rho}_{AB} = \frac{1}{\mu_{\mathcal{G}}(\mathcal{G})} \int_{\mathcal{G}} d\mu_{\mathcal{G}}(g)\, g(\rho_{AB}) \quad \text{(F3)}$$

for any $\rho_{AB} \in (\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n}$ and where $g(\rho_{AB})$ is the image of the mode state ρ_{AB} by g.

If we consider here for $\mathcal{G}$ the group of conjugate passive symplectic operations applied on Alice's n modes and Bob's n modes (in phase space, such operations are simply conjugate orthogonal transformations), then for a given operation g applied to the state, the map $\mathcal{K}_g$ is obtained by applying the orthogonal transformations corresponding to g on the classical data measured by Alice and Bob. If the protocol is invariant under this whole group, then it is sufficient to look at Gaussian states to prove security against collective attacks [6]. In the prepare-and-measure version of the protocol, this group becomes the orthogonal group $O(2n)$.

One simple way to ensure that a protocol is indeed invariant under a set $\mathcal{G}$ of transformations is for Alice and Bob to actively apply random transformations of $\mathcal{G}$ to their states. In particular, if Alice and Bob both apply random orthogonal transformations to their classical vectors in the prepare-and-measure protocol, then the security analysis can be done assuming that they share a Gaussian state in the entangled version of the protocol.

For our security proof, the goal of the symmetrization is to make sure that the state shared by Alice and Bob is as isotropic as possible. Indeed, remember that we need to estimate the covariance matrix of the state shared by Alice and Bob after Alice's generalized measurement $\{\Pi_d, \mathbb{1} - \Pi_d\}$. The only problem that could potentially happen would be that Eve guesses which states might be used for key distillation and which ones might be used for parameter estimation and that she manages to somehow play with Alice and Bob's correlations in order to fool them into overestimating their correlations. To totally prevent such a (quite unrealistic) scenario, it is sufficient to symmetrize the state so that there are no privileged directions in phase space that Eve could exploit. Hence, in practice, the symmetrization does not require one to apply random conjugate passive symplectic operations chosen with the uniform measure over the whole set of such operations: a smaller subset should be efficient. A very conservative quantitative criterium to evaluate the quality of such a set would be, for instance, the distance between the partial trace of $\bar{\rho}_{AB}$ once we trace out $n-1$ modes and the Gaussian state with the same first two moments.

1. *Active symmetrization*

To make sure that the protocol is indeed invariant under specific transformations, we apply an active symmetrization step to the state ρ_{AB} before applying the protocol. The transformations we consider are conjugate passive symplectic operations applied on Alice's n modes and Bob's n modes, which therefore correspond to orthogonal transformations applied to their classical vectors in the prepare-and-measure protocol. For simplicity, we restrict the discussion to this prepare-and-measure scheme in the following.

The active symmetrization requires us to choose a subset $\mathcal{F}$ of the orthogonal group and for Alice and Bob, and to apply the same element $f \in \mathcal{F}$ (chosen uniformly at random) to their data before starting the postprocessing. As we stressed above, taking for $\mathcal{F}$ the whole orthogonal group is not necessary in practice. Hence, we want $\mathcal{F}$ to be a subset of the orthogonal group with the following properties: drawing a random element f from the uniform measure on $\mathcal{F}$ should be doable with resources (time and Alea generation) scaling at most linearly in n, the description of f should also be at most linear in n, and applying f (or f^{-1}) to a random vector of $\mathbb{R}^n$ should also be at most linear in n. These conditions ensure that the protocol with the active symmetrization remains practical. Moreover, the symmetrization should work as well as possible, meaning that $\mathcal{F}$ should symmetrize the state as much as possible. We give examples of such possible subsets $\mathcal{F}$ in the next subsection.

2. *Construction of practical symmetrizations*

Let us describe a recursive algorithm that allows one to draw an orthogonal transformation with the Haar measure on $O(n)$. If we assume that we already drew a random transformation $\tilde{R}_{n-1}$ from the Haar measure on $O(n-1)$, then let us note $R_{n-1} = \mathbb{1} \oplus \tilde{R}_{n-1}$ the orthogonal transformation in

$O(n)$ acting as the identity on the first element of the canonical basis of $\mathbb{R}^n$ and as $\tilde{R}_{n-1}$ on the last $n-1$ elements of the basis. Then, let us draw uniformly at random a unit vector u_n on the sphere $\mathcal{S}^{n-1}$ and define the following Householder reflection H_n:

$$H_n = \mathbb{1}_n - 2u_n u_n^T. \tag{F4}$$

Note that drawing u_n can be done in linear time simply by drawing n random normal variables and normalizing the obtained vector. Also, applying H_n to any vector can be done in linear time in n.

Finally, one can show that random orthogonal transformation $R_n = H_n R_{n-1}$ follows the Haar distribution of $O(n)$ [32]. In particular, one has

$$R_n = \prod_{k=n}^{1} \tilde{H}_k, \tag{F5}$$

where one defines $\tilde{H}_k = \mathbb{1}_{n-k} \oplus H_k$.

Hence, drawing, describing, and applying a random orthogonal transformation from $O(n)$ are tasks with complexity quadratic in n. We define the subset $\mathcal{F}_k$ of the orthogonal group as corresponding to the set of compositions of k such Householder reflections, that is, the last k steps of the algorithm described above. For instance, $\mathcal{F}_1$ corresponds to the set of Householder reflections with respect to a hyperplane of $\mathbb{R}^n$, and $\mathcal{F}_n$ is the orthogonal group $O(n)$. One can also define a family of measures $\mu_1, \mu_2, \ldots, \mu_n$ on $O(n)$ corresponding to the uniform measures of $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$.

A complete symmetrization would imply performing orthogonal transformations on both Alice and Bob's data chosen randomly with the measure μ_n; but for all practical purposes, it seems that μ_1 already provides a high level of symmetrization.

APPENDIX G: FULL PROTOCOL WITH THE SYMMETRIZATION STEP

We present here two different schemes, depending on the choice of modulation which can be either fully Gaussian or consists of a non-Gaussian modulation supplemented by appropriate decoy states. The former modulation is compatible with both a homodyne or a heterodyne detection and corresponds to the technique detailed in Appendix C1 while the latter, which is more efficient in terms of resources, is only compatible with a heterodyne detection (and is detailed in Appendix C). These schemes include the active symmetrization introduced in Appendix F.

1. Fully Gaussian modulation

The full protocol is the following:

(1) Alice draws $2n$ random variables $x_1, x_2, \ldots, x_{2n}$ from a centered normal distribution with the appropriate variance. These form a vector $x \in \mathbb{R}^{2n}$.

(2) Alice sends the states $|\alpha_1\rangle, \ldots, |\alpha_k\rangle, \ldots, |\alpha_n\rangle$ to Bob, with $|\alpha_k\rangle = |x_{2k} + i x_{2k+1}\rangle$.

(3) Bob receives the states after the quantum channel and measures them, with either a homodyne detection or a heterodyne detection. In the case of a heterodyne detection, he obtains a $2n$-dimensional vector y. In the case of a homodyne detection, he obtains an n-dimensional vector y, then informs Alice about his choices of measured quadratures (x or p for each state); Alice only keeps the relevant coordinates in her data in order to form a new n-dimensional vector x.

(4) Alice randomly draws a random orthogonal transformation R from the orthogonal group $O(2n)$ [or $O(n)$ for a homodyne detection]. In theory, to achieve a perfect symmetrization of the state, Alice should draw R with the Haar measure on $O(2n)$. However, in practice, R can be chosen uniformly in a well-chosen subset of $O(2n)$ which has the advantage of allowing for efficient descriptions of its elements, such as one of the measures μ_k defined in Appendix F.

(5) Alice describes R to Bob through the classical communication channel, and both parties apply R to their respective vector, hence obtaining $x' = Rx$ and $y' = Ry$.

(6) Alice chooses randomly n_{PE} coordinates that are used for parameter estimation.

(7) The next step is where lies the novelty of our protocols. For instance, in the so-called four-state protocol, Alice considers the coordinates x'_k which were not used for parameter estimation and keeps only the ones such that $|x'_k|$ is sufficiently close to a predetermined value. In the case of the eight-dimensional protocol (with heterodyne detection, for instance), Alice divides her data into blocks of size 8 and keeps the blocks for which the Euclidean norm is close to a predetermined value (see Appendix C1 for details). Alice informs Bob of the indices that she keeps. The rest of the data are discarded. At this point, Alice and Bob have classical correlations for which an efficient reconciliation protocol is available (see Appendix B).

2. Non-Gaussian modulation and decoy states

In this scheme, the positions of the states used for parameter estimation are chosen randomly *beforehand* by Alice. Let us consider for simplicity the case of the eight-dimensional protocol.

(1) Alice draws n eight-dimensional random vectors, each chosen from one of the three following distributions: random vectors on the seven-dimensional sphere with the appropriate radius (these data correspond to the non-Gaussian modulation which will be used for the key distillation), random vectors on the seven-dimensional sphere with an appropriately fluctuating radius (these are the decoy states which will be discarded at the end of the protocol: the mixture of these states with the previous ones should be indistinguishable from a true Gaussian distribution), or Gaussian vectors which are used for parameter estimation. Alice hence obtains an $8n$-dimensional vector x for which each subset of length 8 corresponds either to legitimate information, decoy data that will be discarded, or data used for parameter estimation.

(2) Alice randomly draws a random orthogonal transformation R from the orthogonal group $O(8n)$. In theory, to achieve a perfect symmetrization of the state, Alice should draw R with the Haar measure on $O(8n)$. However, in practice, R can be chosen uniformly in a well-chosen subset of $O(2n)$ which has the advantage of allowing for efficient descriptions of its elements, such as one of the measures μ_k defined in Appendix F.

(3) Alice computes the vector $x' = Rx$, which is the image of x by the orthogonal transformation R, and uses

this vector for her modulation. Hence she sends the states $|\alpha_1\rangle, \ldots, |\alpha_k\rangle, \ldots, |\alpha_{4n}\rangle$ to Bob, with $|\alpha_k\rangle = |x'_{2k} + ix'_{2k+1}\rangle$.

(4) Bob receives the states after the quantum channel and measures them, with a heterodyne detection. He obtains an $8n$-dimensional vector y'.

(5) Alice describes R to Bob through the classical communication channel. Bob applies R^{-1} to his vector y' and obtains $y = R^{-1}y'$.

(6) Alice reveals which subsets of length 8 should be kept for the key distillation, which ones should be discarded (as they correspond to decoy states), and which ones should be used for parameter estimation.

(7) At this point, Alice and Bob have classical correlations for which an efficient reconciliation protocol is available (see Appendix B).

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

[2] N. J. Cerf and P. Grangier, J. Opt. Soc. Am. B **24**, 324 (2007).

[3] F. Grosshans and P. Grangier, Phys. Rev. Lett. **88**, 057902 (2002).

[4] R. García-Patrón and N. J. Cerf, Phys. Rev. Lett. **97**, 190503 (2006).

[5] M. Navascués, F. Grosshans, and A. Acín, Phys. Rev. Lett. **97**, 190502 (2006).

[6] A. Leverrier and P. Grangier, Phys. Rev. A **81**, 062314 (2010).

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

[8] J. Lodewyck *et al.*, Phys. Rev. A **76**, 042305 (2007).

[9] S. Fossier, E. Diamanti, T. Debuisschert, A. Villing, R. Tualle-Brouri, and P. Grangier, New J. Phys. **11**, 045023 (2009).

[10] A. Leverrier, R. Alléaume, J. Boutros, G. Zémor, and P. Grangier, Phys. Rev. A **77**, 042325 (2008).

[11] M. M. Wolf, G. Giedke, and J. I. Cirac, Phys. Rev. Lett. **96**, 080502 (2006).

[12] A. Leverrier, E. Karpov, P. Grangier, and N. J. Cerf, New J. Phys. **11**, 115009 (2009).

[13] A. Leverrier, F. Grosshans, and P. Grangier, Phys. Rev. A **81**, 062343 (2010).

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

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

[16] A. Leverrier and P. Grangier, Phys. Rev. Lett. **102**, 180504 (2009).

[17] Y.-B. Zhao, M. Heid, J. Rigas, and N. Lütkenhaus, Phys. Rev. A **79**, 012307 (2009).

[18] A. Leverrier and P. Grangier, e-print arXiv:1002.4083.

[19] A. Leverrier and P. Grangier, e-print arXiv:1005.0328.

[20] I. Devetak and A. Winter, Proc. R. Soc. London Ser. A **461**, 207 (2005).

[21] F. Grosshans, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier, Quantum Inf. Comput. **3**, 535 (2003).

[22] V. Scarani and R. Renner, Phys. Rev. Lett. **100**, 200501 (2008).

[23] R. Y. Q. Cai and V. Scarani, New J. Phys. **11**, 045024 (2009).

[24] M. Christandl, R. König, and R. Renner, Phys. Rev. Lett. **102**, 020504 (2009).

[25] L. Sheridan, T. P. Le, and V. Scarani, New J. Phys. **12**, 123019 (2010).

[26] P. Grangier, J. Levenson, and J. Poizat, Nature (London) **396**, 537 (1998).

[27] A. D. Wyner, Bell Syst. Tech. J. **54**, 1355 (1975).

[28] T. Richardson, M. Shokrollahi, and R. Urbanke, IEEE Trans. Inf. Theory **47**, 619 (2001).

[29] T. Richardson and R. Urbanke, Workshop honoring Prof. Bob McEliece on his 60th birthday, 2002, p. 24 (unpublished).

[30] A. Leverrier, Ph.D. thesis, Ecole Nationale Supérieure des Télécommunications, 2009, [http://tel.archives-ouvertes.fr/tel-00451021].

[31] A. Leverrier and N. J. Cerf, Phys. Rev. A **80**, 010102 (2009).

[32] F. Mezzadri, e-print arXiv:math-ph/0609050.