Unconditional Optimality of Gaussian Attacks against Continuous-Variable Quantum Key Distribution

Raúl García-Patrón and Nicolas J. Cerf

QuIC, Ecole Polytechnique, CP 165, Université Libre de Bruxelles, 1050 Bruxelles, Belgium

(Received 4 August 2006; published 10 November 2006)

A fully general approach to the security analysis of continuous-variable quantum key distribution (CV-QKD) is presented. Provided that the quantum channel is estimated via the covariance matrix of the quadratures, Gaussian attacks are shown to be optimal against all collective eavesdropping strategies. The proof is made strikingly simple by combining a physical model of measurement, an entanglement-based description of CV-QKD, and a recent powerful result on the extremality of Gaussian states [M. M. Wolf *et al.*, Phys. Rev. Lett. **96**, 080502 (2006)].

DOI: 10.1103/PhysRevLett.97.190503 PACS numbers: 03.67.Dd, 42.50.−p, 89.70.+c

Continuous-variable (CV) quantum information [1] has attracted a rapidly increasing interest over the past few years. Several quantum key distribution (QKD) schemes based on a Gaussian modulation of coherent states of light combined with homodyne or heterodyne detection have been proposed [2,3] and experimentally demonstrated [4,5]. These protocols have the advantage of being based on standard optical telecom components and thereby of working at high repetition rates compared to the schemes based on single-photon detectors. The first security proof of CV-QKD was restricted to Gaussian individuals attacks [2–4,6]. In such an attack, the eavesdropper (Eve) is assumed to interact individually—according to a Gaussian map—with each of the signal pulses sent over the line, and then to perform a Gaussian (homodyne or heterodyne) measurement on her probe after the basis information (if any) is disclosed but before the full classical postprocessing. Later on, it was shown that non-Gaussian individual attacks cannot beat Gaussian attacks [7], so that studying the security against Gaussian individual attacks is quite justified. This proof extends to the case where Eve attacks finite-size blocks of pulses, but does not cover the important class of collective attacks, where Eve jointly measures all her probes (each having interacted with a signal pulse) after the classical postprocessing has taken place [8–10]. The security versus Gaussian collective attacks were recently studied in [11,12], but a definitive proof of the optimality of Gaussian attacks was missing.

In this Letter, we prove that the optimal collective attack reduces to a Gaussian attack that is completely characterized by the covariance matrix of the quadratures observed by the emitter (Alice) and receiver (Bob). This optimality is plausibly even stronger in view of the fact that, in discrete-variable QKD, the most general attacks, namely, coherent attacks (where Eve coherently interacts with all signal pulses and performs a joint measurement after the classical postprocessing), cannot outperform collective attacks [8,9], implying that it is sufficient to check the security against collective attacks.

One-way QKD protocols with Gaussian continuous variables are divided in two steps, a quantum communication part followed by a classical postprocessing. In the quantum part, Alice sends either a displaced squeezed state encoding a random Gaussian variable or a displaced coherent state encoding two Gaussian variables. Then, Bob performs either homodyne (active basis choice) or heterodyne measurement (no basis choice) on the received states (not necessarily Gaussian) in order to decode Alice's variable. Once Alice and Bob have collected a sufficiently large list of correlated data, they proceed with the classical postprocessing. Unless Alice sent coherent states and Bob did a heterodyne measurement, they first apply a sifting, where they compare the chosen encoding and measurement quadratures (x or p) and keep only the values for which the quadratures match. Then, they apply parameter estimation; i.e., they calculate the covariance matrix γ_{AB} of their correlated variables from a randomly chosen sample of their data. The optimal attack being Gaussian (as we will prove below), γ_{AB} completely characterizes the channel as the first-order moments of the quadratures do not play any role. Finally, they apply one-way error correction and privacy amplification to distill a secret key. The error correction can be done in two ways: either direct reconciliation (DR), where Bob corrects his data to Alice's ones, or reverse reconciliation (RR), where Alice's and Bob's roles are interchanged [4].

Physical model of measurement.—Assume Alice and Bob share a quantum state ρ_{AB} and Alice then makes a von Neumann measurement on system A, obtaining the outcome a distributed according to the probability distribution $p(a)$. This measurement can be realized by applying an appropriate unitary operation U_A on A together with an ancilla, and subsequently observing the state of this ancilla while tracing over the resulting quantum system A' (see Fig. 1). Considering the ancilla as a physical system, noted as a after the action of U_A, the joint state of a and B after the measurement is

$$\rho_{aB} = \int da\, p(a)|a\rangle\langle a| \otimes \rho_B^a. \tag{1}$$

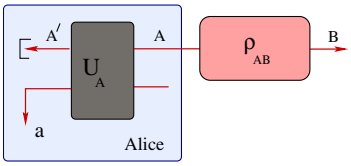

FIG. 1 (color online). Alice's measurement of system A of the bipartite state ρ_{AB}, giving the result a. Equivalently, a denotes the internal state of a preparer who prepares system B according to a.

Given the block-diagonal structure of ρ_{aB}, the quantum mutual entropy $S(a{:}B)$ can be shown to coincide with the Holevo bound $\chi_{aB} = S(\rho_B) - \int dap(a)S(\rho_B^a)$ [13]. Note that the situation here is fully equivalent to that where a is a classical preparer and B is a quantum preparation. Now, assume Bob measures his system B by means of the unitary U_B in a similar way as Alice. The resulting joint state is given by the diagonal density operator,

$$\rho_{ab} = \int dadbp(a,b)|a\rangle\langle a| \otimes |b\rangle\langle b|. \tag{2}$$

The quantum mutual entropy $S(a{:}b)$ then simply reduces to the Shannon mutual information I_{ab} between the preparer's and the measurer's internal states. The Holevo bound on the accessible information then becomes a straightforward consequence of the strong subadditivity of von Neumann entropies, namely [13],

$$I_{ab} = S(a{:}b) \leq S(a{:}bB') = S(a{:}B) = \chi_{aB}. \tag{3}$$

Entanglement-based version of CV-QKD.—The description of any prepare-and-measure CV-QKD protocol using its equivalent entanglement-based scheme is very convenient for security analyses [14]. Indeed, all protocols based on the Gaussian modulation of Gaussian states and homodyne (or heterodyne) measurement can be described in a unified way; see Fig. 2. Alice and Bob are assumed to share a bipartite quantum state ρ_{AB}, whose purification is given to Eve. Alice's measurement of A is equivalent to a preparation scheme where she randomly chooses a, according to $p(a)$, and sends the state $\rho_{B_0}^a$ in the quantum channel so that Bob receives the state ρ_B^a at the output. The unitary U_A determines which measurement is performed: homodyne measurements, corresponding to the preparation of squeezed states, or heterodyne measurements, corresponding to the preparation of coherent states (a then collectively denotes two real numbers). The maximal information that is accessible to Bob is given, in principle, by $\chi_{aB} = S(a{:}B)$. In practice, however, Bob applies an homodyne (or heterodyne) measurement on B, giving b, so the actually extracted information is $I_{ab} = S(a{:}b)$. Since there are two possible encodings at Alice's station and two possible measurements at Bob's station, there exist four Gaussian protocols (three of them having been described in [2,3,6]).

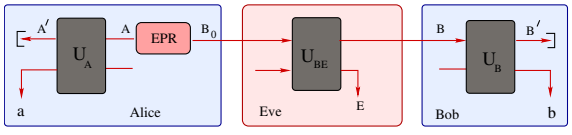

FIG. 2 (color online). Entanglement-based scheme for CV-QKD. Alice's preparation is modeled by a measurement U_A on her half of an EPR pair. The channel is modeled by an unitary interaction between mode B and Eve ancilla's E. Finally, Bob's measurement is modeled by U_B.

Consider now that Eve performs a collective attack: she interacts individually with each signal pulse sent by Alice, stores her resulting probes in a quantum memory, and then applies a joint measurement on them at the end of the classical postprocessing. As shown in [8,9], her information is then limited by the Holevo bound $\chi_{aE} = S(\rho_E) - \int dap(a)S(\rho_E^a)$. Because Eve holds the purification of ρ_{AB}, this bound can be calculated from ρ_{AB}: for example, when Alice and Bob apply the same measurement, it reads $\chi_{aE} = S(\rho_{AB}) - \int dap(a)S(\rho_B^a)$. If ρ_{AB} is assumed to be Gaussian, then χ_{aE} can be directly computed from γ_{AB} [11,12].

Extremality of Gaussian states.—To prove the optimality of Gaussian collective attacks, we also need a very useful theorem, recently proven in [15]. Let us sketch it here for bipartite states ρ_{AB} that have zero first-order moments. Let f be a function satisfying the following properties.

(1) Continuity in trace norm: If $\|\rho_{AB}^{(n)} - \rho_{AB}\|_1 \to 0$ when $n \to \infty$, then $f(\rho_{AB}^{(n)}) \to f(\rho_{AB})$.

(2) Invariance under local "Gaussification" unitaries: $f(U_G^\dagger \otimes U_G^\dagger \rho_{AB}^{\otimes N} U_G \otimes U_G) = f(\rho_{AB}^{\otimes N})$.

(3) Strong superadditivity: $f(\rho_{A_{1,\ldots,N}B_{1,\ldots,N}}) \geq f(\rho_{A_1B_1}) + \cdots + f(\rho_{A_NB_N})$ with equality if $\rho_{A_{1,\ldots,N}B_{1,\ldots,N}} = \rho_{A_1B_1} \otimes \cdots \otimes \rho_{A_NB_N}$.

Then, for every bipartite state ρ_{AB} with covariance matrix γ_{AB}, we have that

$$f(\rho_{AB}) \geq f(\rho_{AB}^G), \tag{4}$$

where ρ_{AB}^G is the Gaussian state with the same γ_{AB}. The proof can be summarized by

$$\begin{aligned} f(\rho_{AB}) &\overset{3}{=} \frac{1}{N} f(\rho_{AB}^{\otimes N}) \overset{2}{=} \frac{1}{N} f(\tilde{\rho}_{A_{1,\ldots,N}B_{1,\ldots,N}}) \\ &\overset{3}{\geq} \frac{1}{N} \sum_{k=1}^{N} f(\tilde{\rho}_{A_kB_k}) \overset{1,\star}{\cong} f(\rho_{AB}^G), \end{aligned} \tag{5}$$

where the superscripts label the assumptions used in each step, while $\tilde{\rho}_{A_{1,\ldots,N}B_{1,\ldots,N}} \equiv U_G^\dagger \otimes U_G^\dagger \rho_{AB}^{\otimes N} U_G \otimes U_G$. The $\star$ stands for the use of a central limit result for quantum states (see [15] for details). The Gaussification unitary U_G is a passive operation, which can be realized with a network of beam splitters and phase shifters. Importantly for what follows, the x and p quadratures of all N modes are not mixed via Gaussification.

Optimality of Gaussian attacks.—The core of our proof now consists in combining this extremality result with the entanglement-based version of CV-QKD supplemented with our physical model of measurement. In realistic protocols, Alice and Bob do not achieve the Holevo bound, but only extract the mutual information $I_{ab} = S(a{:}b)$. In contrast, Eve is assumed to have no technological limitation, so, by collective attacks, she can attain the Holevo bound $\chi_{aE} = S(a{:}E)$. Then, using our notation, the achievable DR secret key rate reads [8,9]

$$K(\rho_{AB}) = S(a{:}b) - S(a{:}E) = S(a|E) - S(a|b). \quad (6)$$

The function $K(\rho_{AB})$ depends on the choice of the measurement done by Alice and Bob (and on the sifting if any), but does not depend on the purification of ρ_{AB}. We now will prove that $K(\rho_{AB})$ satisfies the three conditions of the Gaussian extremality theorem. For this, we also need to use the extension of this function over $2N$ modes ($\bar{A} = A_{1,\ldots,N}$, $\bar{B} = B_{1,\ldots,N}$), namely,

$$K(\rho_{\bar{A}\bar{B}}) = S(\bar{a}{:}\bar{b}) - S(\bar{a}{:}E) = S(\bar{a}|E) - S(\bar{a}|\bar{b}), \quad (7)$$

where Alice (Bob) do the same measurement on her (his) N modes, and Eve has the purification of $\rho_{\bar{A}\bar{B}}$. Note that Eq. (7) restricts to Eq. (6) when $N = 1$.

(i) Continuity: If $\|\rho^{(n)}_{\bar{A}\bar{B}} - \rho_{\bar{A}\bar{B}}\|_1 \leq \epsilon$, using Ulhmann's theorem and the well-known relations between the fidelity and trace distance [16], we can find a purification $|\Psi\rangle^{(n)}_{\bar{A}\bar{B}E}$ ($|\Psi\rangle_{\bar{A}\bar{B}E}$) of $\rho^{(n)}_{\bar{A}\bar{B}}$ ($\rho_{\bar{A}\bar{B}}$) such that $\|\hat{\Psi}^{(n)}_{\bar{A}\bar{B}E} - \hat{\Psi}_{\bar{A}\bar{B}E}\|_1 \leq 2\sqrt{\epsilon}$. Then, considering that partial trace can only decrease the trace norm [16], we have $\|\rho^{(n)}_{\bar{a}E} - \rho_{\bar{a}E}\|_1 \leq 2\sqrt{\epsilon}$ and $\|\rho^{(n)}_{\bar{a}\bar{b}} - \rho_{\bar{a}\bar{b}}\|_1 \leq 2\sqrt{\epsilon}$. Finally, the continuity of von Neumann entropies implies the continuity of K. □

(ii) Invariance under local Gaussification unitaries: Applying the local Gaussification operation $U_G \otimes U_G$ on the product states $|\psi\rangle^{\otimes N}_{ABE}$ (as shown in Fig. 3 for $N = 2$), we obtain the state $|\tilde{\psi}\rangle_{\bar{A}\bar{B}\bar{E}}$. After the measurements on Alice's and Bob's sides, the state becomes $\tilde{\rho}_{\bar{a}\bar{b}\bar{E}}$. But because the (homodyne or heterodyne) measurement and the Gaussification operation can be interchanged, by applying $U^\dagger_G \otimes U^\dagger_G$ on modes $\bar{a}$ and $\bar{b}$ we recover the state $\rho^{\otimes N}_{abE}$, which coincides with the state obtained by directly measuring $|\psi\rangle^{\otimes N}_{ABE}$ without Gaussification. Since the two states $\tilde{\rho}_{\bar{a}\bar{b}}$ and $\rho^{\otimes N}_{ab}$ are related by a local unitary operation $U^\dagger_G \otimes U^\dagger_G$, and since the mutual von Neumann entropies

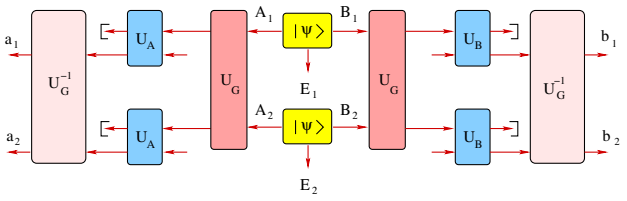

FIG. 3 (color online). Invariance under local Gaussification unitaries: U_G can be interchanged with the measurement U_A; then U_G^{-1} and U_G cancel each other.

appearing in $K(\rho_{AB})$ are invariant under (any) local unitaries, we obtain the invariance of $K(\rho_{AB})$ under local Gaussification unitaries. □

(iii) Strong superadditivity: We will restrict the proof to two modes on each side, $A_{1,2}$ and $B_{1,2}$, the generalization to $N > 2$ being straightforward. We have

$$K(\rho_{A_{1,2}B_{1,2}}) = S(a_1a_2|E) - S(a_1a_2|b_1b_2), \quad (8)$$

where the conditional entropies can be expressed as

$$S(a_1a_2|E) = S(a_1|a_2E) + S(a_2|a_1E) + S(a_1{:}a_2|E),$$
$$S(a_1a_2|b_1b_2) = S(a_1|b_1b_2) + S(a_2|b_1b_2) - S(a_1{:}a_2|b_1b_2).$$

As a consequence of the strong subadditivity of von Neumann entropies, we obtain the bound

$$K \geq \underbrace{S(a_1|a_2E) - S(a_1|b_1b_2)}_{\geq S(a_1|A_2B_2E) - S(a_1|b_1)} + \underbrace{S(a_2|a_1E) - S(a_2|b_1b_2)}_{\geq S(a_2|A_1B_1E) - S(a_2|b_2)} \quad (9)$$

(using the fact that conditioning can only decrease the conditional entropy). The purification of A_1B_1 (A_2B_2) being A_2B_2E (A_1B_1E), we obtain

$$K(\rho_{A_{1,2}B_{1,2}}) \geq K(\rho_{A_1B_1}) + K(\rho_{A_2B_2}). \quad (10)$$

The additivity of $K(\rho_{A_{1,2}B_{1,2}})$ is a straightforward consequence of the additivity of von Neumann entropies. □

Thus, using Eq. (4), we have proved that for all bipartite quantum states ρ_{AB} with covariance matrix γ_{AB}, one has $K(\rho_{AB}) \geq K(\rho^G_{AB})$. This means that $K(\rho^G_{AB})$ is a lower bound on the secret key rate for any protocol (even non-Gaussian) and collective attack (including non-Gaussian). The only requirement for this result to hold is that Alice and Bob use the second-order moments of the quadratures in order to calculate this bound. In particular, for the Gaussian-modulation protocols of [2–4,6], Eve's optimal attack is a Gaussian attack, in which case the bound is saturated. Note that the above proof concerns DR [see Eq. (6)], but its extension to RR is straightforward: one simply needs to interchange $a \leftrightarrow b$ and $A \leftrightarrow B$. Importantly, this bound can easily be computed from the observed data since one simply needs to calculate the entropy of thermal states. As an illustration, Fig. 4 shows the security range of Gaussian-modulation protocols against Gaussian collective attacks.

Coherent attacks.—They represent the most powerful class of attacks Eve can perform: she let all the signal pulses sent by Alice interact with a large auxiliary system (quantum computer), which she measures jointly at the end of the classical postprocessing. Recently, it has been shown that, for discrete-variable QKD and under some symmetries of the classical postprocessing, collective attacks are actually as efficient for Eve as coherent attacks [8,9]. Taking for granted that this result extends to CV-QKD, we conjecture that our optimality proof of Gaussian attacks holds in full generality.

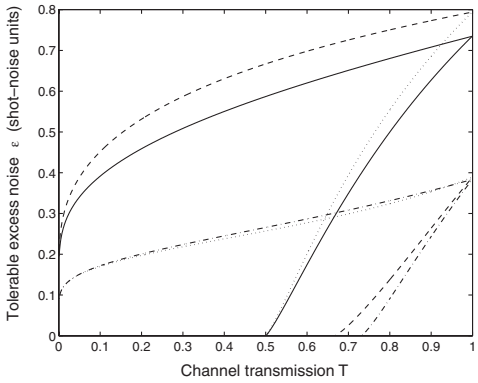

FIG. 4. Tolerable excess noise ϵ as a function of the channel transmission T at the limit of an infinite modulation for the four Gaussian protocols: squeezed states and homodyne measurement (solid line; see also [12]), squeezed states and heterodyne measurement (dashed line), coherent states and homodyne measurement (dotted line; see also [12]), and coherent states and heterodyne measurement (dot-dashed line). The curves vanishing at (or above) $T = 0.5$ correspond to DR, whereas those vanishing at $T = 0$ refer to RR.

Realistic implementations of CV-QKD.—They never achieve the secret key rate $K(\rho_{AB})$ because reconciliation protocols are not 100% efficient. The actual key rate is

$$K = \beta S(a{:}b) - S(a{:}E) = S(a|E) - \beta S(a|b) - (1-\beta)S(a), \quad (11)$$

where $\beta \in [0, 1]$ is the reconciliation efficiency. It is easy to prove that Eq. (11) also satisfies the three conditions of the extremality theorem, so our conclusions remain unchanged. In the special case of $\beta = 0$, this means that Eve's accessible information $\chi_{aE} = S(a{:}E)$ is maximized for Gaussian states, so that Gaussian collective attacks are also optimal in this restricted sense.

"Quantum" Bob.—A theoretically interesting, though probably unrealistic, situation is the case where Bob reaches the Holevo bound χ_{aB}. This may be done by combining the use of quantum memory with a proper optimal postprocessing at Bob's side. The "ultimate" available secret key rate then reads

$$K = S(a{:}B) - S(a{:}E) = S(a|E) - S(a|B). \quad (12)$$

It again satisfies the three above conditions, so it is lower bounded by the Gaussian attack.

Conclusion.—We have presented a unified analysis of all known QKD protocols based on Gaussian modulation of coherent (or squeezed) states by Alice and homodyne (or heterodyne) detection by Bob, for the DR and RR versions of one-way reconciliation. The entanglement-based model of CV-QKD combined with a physical representation of measurement gives a very simple way of writing the secret key rates in terms of mutual von Neumann entropies involving quantum systems (including the preparer and the measurer). Then, exploiting a recent result on the extremality of Gaussian states, we have demonstrated that the optimal collective attack against all these protocols is a Gaussian operation. It is then sufficient to check the security against Gaussian attacks, which are completely characterized by the covariance matrix γ_{AB} estimated by Alice and Bob. This result appears to be quite general as it holds for realistic protocols (with finite reconciliation efficiency) as well as for ideal protocols (where Bob has a quantum memory and extracts the entire accessible information). Provided that [8,9] can be adapted to CV, which is a topic for further investigation, our proof would extend to the full unconditional security of CV-QKD against coherent attacks.

We acknowledge financial support from the EU under projects COVAQIAL (FP6-511004) and SECOQC (IST-2002-506813) and from the IUAP programme of the Belgian government under Grant No. V-18. R. G.-P. acknowledges support from the Belgian foundation FRIA.

Note added.—The optimality of Gaussian collective attacks has been independently proved using different techniques in [17].

[1] S. L. Braunstein and A. K. Pati, *Quantum Information Theory with Continuous Variables* (Kluwer Academic, Dordrecht, 2003).
[2] F. Grosshans and P. Grangier, Phys. Rev. Lett. **88**, 057902 (2002).
[3] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam, Phys. Rev. Lett. **93**, 170504 (2004).
[4] F. Grosshans, G. Van Assche, J. Wenger, R. Tualle-Brouri, N. J. Cerf, and P. Grangier, Nature (London) **421**, 238 (2003).
[5] J. Lodewyck, T. Debuisschert, R. Tualle-Brouri, and P. Grangier, Phys. Rev. A **72**, 050303(R) (2005).
[6] N. J. Cerf, M. Lévy, and G. Van Assche, Phys. Rev. A **63**, 052311 (2001).
[7] F. Grosshans and N. J. Cerf, Phys. Rev. Lett. **92**, 047905 (2004).
[8] R. Renner, N. Gisin, and B. Kraus, Phys. Rev. A **72**, 012332 (2005).
[9] R. Renner, Ph.D. thesis, ETH Zürich, 2005.
[10] I. Devetak and A. Winter, Phys. Rev. Lett. **93**, 080501 (2004).
[11] F. Grosshans, Phys. Rev. Lett. **94**, 020504 (2005).
[12] M. Navascues and A. Acin, Phys. Rev. Lett. **94**, 020505 (2005).
[13] N. J. Cerf and C. Adami, quant-ph/9611032.
[14] F. Grosshans, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier, Quantum Inf. Comput. **3**, 535 (2003).
[15] M. M. Wolf, G. Giedke, and J. I. Cirac, Phys. Rev. Lett. **96**, 080502 (2006).
[16] M. A. Nielsen and I. C. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2002).
[17] M. Navascués, F. Grosshans, and A. Acín, Phys. Rev. Lett. **97**, 190502 (2006).