Extremality of Gaussian Quantum States

Michael M. Wolf,[1] Geza Giedke,[1,2] and J. Ignacio Cirac[1]

[1]*Max-Planck-Institute for Quantum Optics, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*
[2]*Institut für Quantenelektronik, ETH Zürich, Wolfgang-Pauli-Strasse 16, CH-8093 Zürich, Switzerland*

(Received 27 September 2005; published 2 March 2006)

We investigate Gaussian quantum states in view of their exceptional role within the space of all continuous variables states. A general method for deriving extremality results is provided and applied to entanglement measures, secret key distillation and the classical capacity of bosonic quantum channels. We prove that for every given covariance matrix the distillable secret key rate and the entanglement, if measured appropriately, are minimized by Gaussian states. This result leads to a clearer picture of the validity of frequently made Gaussian approximations. Moreover, it implies that Gaussian encodings are optimal for the transmission of classical information through bosonic channels, if the capacity is additive.

DOI: 10.1103/PhysRevLett.96.080502 PACS numbers: 03.67.Hk, 03.65.Ud

States with a Gaussian Wigner distribution, so-called *Gaussian states*, appear naturally in every quantum system which can be described or approximated by a quadratic bosonic Hamiltonian. They are ubiquitous in quantum optics as well as in the description of atomic ensembles, ion traps or nanomechanical oscillators. Moreover, Gaussian states became the core of quantum information theory with continuous variables.

Besides their practical relevance, Gaussian states play an exceptional role with respect to many of their theoretical properties. A particular property of Gaussian states is that they tend to be extremal within *all* continuous variable states if one imposes constraints on the covariance matrix (CM). The best known example of that kind is the extremality with respect to the entropy: within all states having a given CM, Gaussian states attain the maximum von Neumann entropy (cf. [1]). Similar extremality properties have recently been shown for the mutual information [2] and conditional entropies [3,4].

In this Letter we prove extremality results for Gaussian states with respect to entanglement measures, secret key rates, and the classical capacity of bosonic quantum channels. These findings are based on a general method, which exploits the central limit theorem as a active and local *Gaussification* operation. Our main focus lies on the entanglement, which will serve as a showcase for the general procedure. We prove that for any given CM, the entanglement, if measured in an appropriate way, is lower bounded by that of a Gaussian state. The same result is shown to hold true for many other quantities like the distillable randomness and the secret key rate. This result not only emphasizes the exceptional role of Gaussian states, it also leads to a clearer picture of the validity of frequently made Gaussian approximations. In practice, states deviate from exact Gaussians and their precise nature remains mostly unknown. Nevertheless, the CM can typically be determined, e.g., by homodyne detection, and one is tempted to calculate the amount of entanglement, or other quantities, from the CM under the assumption that the state is Gaussian. The derived extremality of Gaussian states now justifies this approach as it excludes an overestimation of the desired quantity, even in cases where the actual state is highly non-Gaussian. In this sense one stays on the safe side when *a priori* assuming the state to be Gaussian. We will see, however, that some care is in order, since the extremality property with respect to the entanglement turns out to depend on the chosen entanglement measure.

Before we derive the main results we will briefly recall the basic notions. Consider a bosonic system of N modes characterized by N pairs of canonical operators $(Q_1, P_1, \ldots, Q_N, P_N) =: R$ or equivalently by N bosonic annihilation operators $a_j = (Q_j + iP_j)/\sqrt{2}$. For any density operator ρ of the system we define a vector of means with components $d_j = \mathrm{tr}[\rho R_j]$, a CM $\Gamma_{kl} = \mathrm{tr}[\rho\{R_k - d_k, R_l - d_l\}_+]$ and introduce a *characteristic function* $\chi(\xi) = \mathrm{tr}[\rho \exp(i\xi \cdot R)]$, $\xi \in \mathbb{R}^{2N}$. The latter is the Fourier transform of the Wigner function and it thus completely characterizes the state [5]. For Gaussian states, the characteristic function has the form

$$\chi(\xi) = e^{i\xi\cdot d - 1/4\xi\cdot\Gamma\xi}, \tag{1}$$

such that they are entirely specified by d and Γ leading to a complete description within a finite dimensional phase space $\mathbb{R}^{2N}$. The underlying Hilbert space $\mathcal{H}$ is, however, infinite dimensional and we will denote the set of all bounded linear operators on $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$. Note that the following results also apply to finite dimensional systems by simply embedding $\mathbb{C}^d$ into $\mathcal{H}$.

Our results are based on a noncommutative central limit theorem, as discussed in [6,7], and operator-topology arguments from [7,8]. We will first state the main ingredient as a general Lemma and then discuss its applications to quantum information theory. Readers who are mainly interested in the applications may skip the proof of the Lemma.

Lemma 1.—Let f: $\mathcal{B}(\mathcal{H}^{\otimes N}) \to \mathbb{R}$ be a continuous functional, which is strongly superadditive and invariant under

local unitaries $f(U^{\otimes N}\rho U^{\dagger\otimes N}) = f(\rho)$. Then for every density operator ρ describing an N-partite system with finite first and second moments, we have that

$$f(\rho) \geq f(\rho_G), \tag{2}$$

where ρ_G is the Gaussian states with the same first and second moments as ρ.

Let us first remark on the requirements in Lemma 1. *Continuity* is understood in trace norm, i.e., $\|\rho^{(n)} - \rho\|_1 \to 0$ should imply $f(\rho^{(n)}) \to f(\rho)$. In fact, lower semicontinuity suffices, and we may restrict the domain of f to an appropriate compact subset of density operators like those satisfying an energy constraint. The latter typically restores continuity for functionals, which are known to be continuous in the finite dimensional case. *Strong superadditivity* means that given a state ρ acting on $(\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}) \otimes (\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2})$ with restrictions ρ_i to $\mathcal{H}_{A_i} \otimes \mathcal{H}_{B_i}$, then $f(\rho) \geq f(\rho_1) + f(\rho_2)$, with equality if $\rho = \rho_1 \otimes \rho_2$. The latter is referred to as *additivity* and both properties are analogously defined for more than two parties.

Proof.—The main idea of the proof is covered by the following equation:

$$f(\rho) = \frac{1}{n} f(\rho^{\otimes n}) = \frac{1}{n} f(\tilde{\rho}) \tag{3}$$

$$\geq \frac{1}{n} \sum_{k=1}^{n} f(\tilde{\rho}_k) \to f(\rho_G). \tag{4}$$

In the first line we use additivity of f and set $\tilde{\rho} = U^{\otimes N}\rho^{\otimes n}U^{\dagger\otimes N}$, where U is a suitably chosen local unitary, which acts on n copies of the state. In the second line we first exploit strong superadditivity in order to bound $f(\tilde{\rho})$ from below by the sum over all reduced density operators $\tilde{\rho}_k$. Then we argue by the central limit theorem and a special choice of U, that each of these reduced states $\tilde{\rho}_k$ converges to the Gaussian ρ_G in the limit $n \to \infty$.

This idea is now made rigorous in two steps. First, we prove that each characteristic function $\tilde{\chi}_k$ converges pointwise to the corresponding Gaussian χ_G, and then we argue that this implies trace-norm convergence on the level of density operators. To simplify matters we will, without loss of generality, assume that ρ has vanishing first moments, i.e., $d_j = 0$. The general case is then obtained by applying local displacements, which by assumption will not change the value of f.

Let us begin with specifying the local unitary U as a passive symplectic operation acting on the canonical operators on site $\alpha \in \{1, \ldots, N\}$ as

$$\tilde{Q}_{\alpha,k} = \sum_{l=1}^{n} \frac{H_{kl}}{\sqrt{n}} Q_{\alpha,l}, \qquad \tilde{P}_{\alpha,k} = \sum_{l=1}^{n} \frac{H_{kl}}{\sqrt{n}} P_{\alpha,l}, \tag{5}$$

with

$$H = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

being a Hadamard matrix and $n = 2^m$. Physically, U corresponds to an array of 50:50 beam splitters and half wave plates. Note that H has only entries ± 1, so that we can partition the sum

$$\tilde{Q}_{\alpha,k} = \frac{1}{\sqrt{n}} \Big(\sum_{l:H_{kl}=1}^{n_+} Q_{\alpha,l} - \sum_{j:H_{kj}=-1}^{n_-} Q_{\alpha,j} \Big), \tag{6}$$

and similarly for $\tilde{P}_{\alpha,k}$. Here $n_+ = n - n_-$ is the number of ones in the kth row of H. Note that either $n_+ = n$ in the first row, or $n_+ = n/2$ in all other rows.

The characteristic function $\tilde{\chi}_k$ of the reduced density operator $\tilde{\rho}_k$ is then given by

$$\tilde{\chi}_k(q, p) = \mathrm{tr}\Big[\rho^{\otimes n} \exp\Big(i \sum_{\alpha=1}^{N} q_\alpha \tilde{Q}_{\alpha,k} + p_\alpha \tilde{P}_{\alpha,k} \Big) \Big] \tag{7}$$

$$= \chi\Big(\frac{\xi}{\sqrt{n}}\Big)^{n_+} \chi\Big(\frac{-\xi}{\sqrt{n}}\Big)^{n_-}, \tag{8}$$

where χ is the characteristic function of ρ and $\xi = (q, p)$.

Following [7] we introduce a function $g: \mathbb{R} \to \mathbb{C}$ by $g(x) = \chi(x\xi)$, which is a *classical* characteristic function, i.e., the Fourier transform of a classical probability distribution with second moment $\xi^T\Gamma\xi/2$. To see this, note that χ is the Fourier transform of the Wigner function and recall that every one-dimensional marginal of a Wigner function (in particular the one corresponding to the direction ξ) is an admissible probability distribution. Characteristic functions are continuous at the origin, satisfy $g(0) = 1$, $|g(x)| \leq 1$, and in the case of finite second moments we can expand up to second order [9], such that

$$g(x) = 1 - \frac{\xi^T\Gamma\xi}{4} x^2 + o(x^2). \tag{9}$$

Pointwise convergence $\tilde{\chi}_k \to \chi_G$ follows then from Eq. (8) together with setting $x = 1$ in

$$\lim_{n\to\infty} g\Big(\frac{x}{\sqrt{n}}\Big)^{n_+} g\Big(\frac{-x}{\sqrt{n}}\Big)^{n_-} = \lim_{n\to\infty} \Big(1 - \frac{\xi^T\Gamma\xi}{4n} x^2\Big)^n \tag{10}$$

$$= \exp\Big[-\frac{1}{4} \xi^T\Gamma\xi x^2 \Big]. \tag{11}$$

For the remaining part we can combine the argumentations in Refs. [7,8]. In [7] it was proven that pointwise convergence of the characteristic functions implies convergence of the respective density operators within the weak operator topology. The latter was, however, shown to be equivalent to the trace-norm topology on density operators in Ref. [8]. □

A simple application of Lemma 1 is the rederivation of the maximum entropy principle by setting f equal to minus the von Neumann entropy $S(\rho) = -\mathrm{tr}[\rho \log \rho]$. Similarly, in the bipartite case with $N = N_A + N_B$ and $f(\rho) = S(\rho_A) - S(\rho)$, we recover the recently proven extremality result for the conditional entropy [4] for which strong

superadditivity is an immediate consequence of the strong subadditivity inequality for the entropy.

Entanglement measures.—Grouping the N tensor factors in Lemma 1 into $M \leq N$ parties and exploiting that every entanglement measure is by definition invariant under local unitaries, yields the following.

Proposition 1.—Let E be a continuous entanglement measure which is strongly superadditive. Then, for every density operator ρ describing an M-partite system with finite CM (and arbitrary, finite, number of modes per site), we have that any Gaussian state ρ_G with the same CM provides a lower bound $E(\rho_G) \leq E(\rho)$.

Most of the entanglement theory developed so far is devoted to bipartite systems. Whereas for pure states there is an essentially unique measure of entanglement, the von Neumann entropy of the reduced state, there are various different entanglement measures for mixed states [10]. Two of them have a clear operational meaning: the *distillable entanglement* E_D quantifies the amount of pure state entanglement that can asymptotically be extracted by means of local operations and classical communication (LOCC), and the *entanglement of formation* E_F (its regularized form, the *entanglement cost* E_c) measures the pure state entanglement required in order to prepare the state. Among all other measures, the *logarithmic negativity* E_N is the most popular one, as it is comparatively easy to calculate [11]. Let us now discuss the consequences of Proposition 1 for some entanglement measures:

(i) Distillable entanglement. E_D is additive (due to the asymptotic definition) and strongly superadditive (since restricted protocols lead to smaller rates). It was shown to be continuous in the interior of the set of distillable states [12]. As distillable Gaussian states are always in the interior [13], E_D fulfills all the requirements in the Lemma and analogous reasonings hold true in the multipartite case where ω is replaced by any M-partite target state. Moreover, in the bipartite case Prop. 1 leads immediately to a simple sufficient distillability criterion, since for Gaussian states it is known that $E_D > 0$ is equivalent to a simple inequality for the CM [13].

(ii) Entanglement of formation. Continuity of E_F was proven in [14] for finite and in [15] for infinite dimensional systems with energy constraint. Superadditivity of E_F is a notorious conjecture, which is proven to be equivalent to additivity of E_F and to many other additivity conjectures [16]. These are known for various special cases (cf. [4,16,17]) but remain to be proven in general.

(iii) Squashed entanglement. This is indeed strongly superadditive [18] and its continuity was proven (for finite dimensions) in Ref. [19].

(iv) Logarithmic negativity. E_N is additive, but fails to be strongly superadditive. In fact, E_N does not only fail the requirements for the proposition—Eq. (2) turns out to be false in this case. A simple counterexample is given by the state $|\varphi\rangle = \sqrt{1-\lambda^2}|00\rangle + \lambda|11\rangle$, with $\lambda = \frac{1}{4}$. In this case $E_N(\varphi) \simeq 0.57$, whereas $E_N(\varphi_G) \simeq 0.64$ [20]. Hence, despite the fact that E_N can easily be calculated, it is not a faithful entanglement measure in the sense that a Gaussian approximation of a non-Gaussian state could lead to an overestimated amount of entanglement.

Finally, it is interesting to note that Gaussian states not only give a lower bound, but they also provide an upper bound to the entanglement if only the CM is known. In fact, it is a simple consequence of the maximum entropy property that for a given energy (i.e., $\mathrm{tr}[\Gamma]$ fixed) the entanglement is maximized by a Gaussian state [21].

Secret key distillation.—We will now depart from the discussion of entanglement and see how Lemma 1 can be applied to the distillation of a classical secret key from quantum states under the assumption of collective attacks. Consider the case where two parties share m copies of a quantum state ρ_{AB} and aim at converting these into rm bits of a secret key under local operations and public communication. Allowing for the worst case scenario, in which an eavesdropper is given the entire purifying system of $\rho_{AB} = \mathrm{tr}_E[|\Psi_{ABE}\rangle\langle\Psi_{ABE}|]$, this can be understood as a mapping $\Psi_{ABE}^{\otimes m} \to \sigma^{\otimes rm} \otimes \rho_E$, where the secret bits $\sigma = \frac{1}{2}(|00\rangle \times \langle 00| + |11\rangle\langle 11|)$ are asymptotically uncorrelated with the state ρ_E of the eavesdropper. We call the asymptotic supremum over all achievable rates r the *distillable secret key* $K^{\mathrm{coll}}(\rho_{AB})$ of the state. By the same reasoning as for the distillable entanglement, K^{coll} (together with its multipartite generalizations with $\sigma = \frac{1}{2}(|0\ldots0\rangle\langle 0\ldots0| + |1\ldots1\rangle\langle 1\ldots1|)$) has the properties of being additive and strongly superadditive. Hence, under the assumption of continuity Lemma 1 implies the following.

Proposition 2.— Consider an M-partite system with an arbitrary, finite, number of modes per site. Then for every given finite first and second moments the Gaussian state ρ_G provides a lower bound to the distillable secret key $K^{\mathrm{coll}}(\rho) \geq K^{\mathrm{coll}}(\rho_G)$.

Channel capacities.—Let us finally apply Lemma 1 to the task of transmitting classical information through a bosonic Gaussian quantum channel T [2,4]. The latter may describe optical fibers, harmonic chains, or any other bosonic system for which we can describe the evolution in terms of a quadratic Hamiltonian H acting on system plus environment

$$T(\rho) = \mathrm{tr}_{\mathrm{env}}[V(\rho \otimes |\phi\rangle\langle\phi|_{\mathrm{env}})V^\dagger], \qquad V = e^{iHt}. \quad (12)$$

The classical capacity C of a quantum channel is the asymptotically achievable number of classical bits that can be reliably transmitted from a sender to a receiver per use of the channel. To make this a reasonable concept in the infinite dimensional setting, one imposes an energy constraint to the encoding. That is, any allowed set of input states ρ_i with respective probabilities p_i is such that the average state $\bar{\rho} = \sum_i p_i \rho_i$ satisfies an energy constraint $\bar{\rho} \in \mathcal{K} = \{\rho | \sum_j \mathrm{tr}[(Q_j^2 + P_j^2)\rho] \leq \kappa\}$. Under this constraint the capacity $C(T, \mathcal{K})$ of the channel is

$$C_1(T, \mathcal{K}) = \sup_{\{p_i,\rho_i\}} [S(T(\bar{\rho})) - \sum_i p_i S(T(\rho_i))], \quad (13)$$

$$C(T, \mathcal{K}) = \lim_{n\to\infty} \frac{1}{n} C_1(T^{\otimes n}, \mathcal{K}^{\otimes n}). \quad (14)$$

Consider now a fixed state $\bar{\rho}$ and define $\rho = V(\bar{\rho} \otimes |\phi\rangle \times \langle\phi|_{\text{env}})V^\dagger$. Then we can write

$$C_1(T, \bar{\rho}) = S(T(\bar{\rho})) - E_F(\rho). \quad (15)$$

If the notorious additivity conjecture is true [16], then not only E_F satisfies the requirements of Lemma 1 but also $C_1 = C$, i.e., the supremum over all $\bar{\rho} \in \mathcal{K}$ in Eq. (15) gives the capacity of T. By Proposition 1 together with the maximum entropy property of Gaussian states we have then, however, that $C(T, \bar{\rho}) \leq C(T, \bar{\rho}_G)$ as both terms in Eq. (15) become extremal for the Gaussian state $\bar{\rho}_G$, which has the same CM as $\bar{\rho}$. Since $\bar{\rho} \in \mathcal{K}$ iff $\bar{\rho}_G \in \mathcal{K}$ this shows:

Proposition 3.—Consider a bosonic Gaussian channel acting on a finite number of modes. Then there is an optimal encoding, which achieves the classical capacity with a Gaussian $\bar{\rho}$, if the capacity is additive.

For single mode channels for which the optimal ρ is a symmetric two-mode Gaussian state we then even know an optimal ensemble $\{p_i, \rho_i\}$ (which is continuous in this case). For such two-mode states it has been shown [22] that $E_F(\rho)$ equals the so-called *Gaussian entanglement of formation* [23], which in turn implies that the optimal ensemble consists of coherent states which are distributed in phase space according to an appropriate Gaussian distribution.

Summary and outlook.—We presented a general method which allows to prove extremality of Gaussian quantum states with regard to various applications. This reemphasizes the exceptional role of these states and what is more, it justifies frequently made Gaussian (i.e., quadratic) approximations. As we saw, in particular, the asymptotic nature of many quantities in quantum information theory fits very well to the asymptotic nature of the central limit theorem. Hence, there are certainly many other similar applications. It is, for instance, straightforward to translate some results from the bosonic to the Fermionic world [24]. Moreover, via a state-channel duality one might apply similar techniques to channels and operations instead of states (cf. [6]). In fact, recently, Gaussian operations turned out to be optimal for certain tasks concerning classical teleportation and cloning of coherent states [25].

The authors are grateful to the Benasque Center for Science, where parts of this work were developed. M. M. W. thanks R. F. Werner and D. Perez-Garcia for interesting discussions.

[1] A. S. Holevo, M. Sohma, and O. Hirota, Phys. Rev. A **59**, 1820 (1999).
[2] A. S. Holevo and R. F. Werner, Phys. Rev. A **63**, 032312 (2001).
[3] F. Grosshans and N. J. Cerf, Phys. Rev. Lett. **92**, 047905 (2004).
[4] J. Eisert and M. M. Wolf, quant-ph/0505151.
[5] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
[6] J. Quaegebeur, J. Funct. Anal. **57**, 1 (1984).
[7] C. D. Cushen and R. L. Hudson, J. Appl. Probab. **8**, 454 (1971).
[8] G. F. Dell'Antonio, Commun. Pure Appl. Math. **20**, 413 (1967); E. B. Davies, Commun. Math. Phys. **15**, 277 (1969); E. B. Davies, Commun. Math. Phys. **27**, 309 (1972).
[9] P. A. P. Moran, *An Introduction to Probability Theory* (Claredon, Oxford, 1968).
[10] M. Horodecki, Quantum Inf. Comput. **1** No. 1, 3 (2001).
[11] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002); M. B. Plenio, Phys. Rev. Lett. **95**, 090503 (2005).
[12] G. Vidal, quant-ph/0203107.
[13] G. Giedke, L.-M. Duan, J. I. Cirac, and P. Zoller, Quantum Inf. Comput. **1** No. 3, 79 (2001); L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **84**, 4002 (2000); R. F. Werner and M. M. Wolf, Phys. Rev. Lett. **86**, 3658 (2001).
[14] M. A. Nielsen, Phys. Rev. A **61**, 064301 (2000).
[15] M. E. Shirokov, quant-ph/0411091.
[16] P. W. Shor, Commun. Math. Phys. **246**, 453 (2004); K. Matsumoto, T. Shimono, and A. Winter, Commun. Math. Phys. **246**, 427 (2004); K. M. R. Audenaert and S. L. Braunstein, Commun. Math. Phys. **246**, 443 (2004).
[17] M. M. Wolf and J. Eisert, New J. Phys. **7**, 93 (2005).
[18] M. Christandl and A. Winter, J. Math. Phys. (N.Y.) **45**, 829 (2004).
[19] R. Alicki and M. Fannes, quant-ph/0312081.
[20] A similar example but with fixed *EPR uncertainty* [22] was found by J. Eisert (private communication).
[21] S. J. v. Enk and O. Hirota, Phys. Rev. A **71**, 062322 (2005).
[22] G. Giedke, M. M. Wolf, O. Krueger, R. F. Werner, and J. I. Cirac, Phys. Rev. Lett. **91**, 107901 (2003).
[23] M. M. Wolf, G. Giedke, O. Krueger, R. F. Werner, and J. I. Cirac, Phys. Rev. A **69**, 052320 (2004).
[24] R. L. Hudson, J. Appl. Probab. **10**, 502 (1973).
[25] N. J. Cerf, O. Krueger, P. Navez, R. F. Werner, and M. M. Wolf, Phys. Rev. Lett. **95**, 070501 (2005); K. Hammerer, M. M. Wolf, E. S. Polzik, and J. I. Cirac, Phys. Rev. Lett. **94**, 150503 (2005).