Quantum Capacities of Bosonic Channels

Michael M. Wolf,[1] David Pérez-García,[2] and Geza Giedke[1]

[1]*Max-Planck-Institute for Quantum Optics, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*

[2]*Área de Matemática Aplicada, Universidad Rey Juan Carlos, Calle Tulipan, 28933 Móstoles (Madrid), Spain*

(Received 15 June 2006; published 26 March 2007)

We investigate the capacity of bosonic quantum channels for the transmission of quantum information. We calculate the quantum capacity for a class of Gaussian channels, including channels describing optical fibers with photon losses, by proving that Gaussian encodings are optimal. For arbitrary channels we show that achievable rates can be determined from few measurable parameters by proving that every channel can asymptotically simulate a Gaussian channel which is characterized by second moments of the initial channel. Along the way we provide a complete characterization of degradable Gaussian channels and those arising from teleportation protocols.

DOI: 10.1103/PhysRevLett.98.130501 PACS numbers: 03.67.Hk, 42.81.Dp

One of the aims of quantum information theory [1] is to follow the ideas of Shannon and to establish a theory of information based on the rules of quantum mechanics. A key problem along this way is the calculation of the quantum capacity of noisy quantum channels. That is, the question how much quantum information (measured in number of qubits) can be transmitted coherently through a channel such as a lossy optical fiber, or stored reliably in a quantum memory—the future version of present-day hard drives? Despite substantial progress [2,3] a general computable formula for this capacity, comparable to Shannon's seminal coding theorem for classical information, is not in sight.

In this Letter, we focus on the special case of bosonic channels, in which a collection of bosonic modes is used to transmit (quantum) information. Arguably, this is the practically most important class of channels, since quantum information is almost invariably sent using photons: be it through the ubiquitous optical fibers, in free space, or in the microwave range via superconducting transmission lines (cf. [4]). In addition to the transmission in space, bosonic channels also play a major role for "transmission in time," i.e., in quantum memories. Several of the most advanced light-matter interfaces [5] make use of atomic ensembles to store photonic quantum information in collective atomic degrees of freedom which are in turn well described by bosonic modes. The quantum capacity of the corresponding channel is an adequate figure of merit for such devices.

The Letter has two parts in which we first deal with incomplete knowledge of the physical channel and then explicitly determine the capacity for some of the most relevant cases. In the first part we prove that the quantum capacity of any bosonic channel T is lower bounded by that of a corresponding Gaussian channel T_G, which can be derived from measurable moments of T. This implies that for determining and certifying achievable rates for the transmission of quantum information through T we need not know the channel exactly (which might be hardly possible in infinite dimensions), but merely its second moments, i.e., a few measurable parameters. In the second part we then explicitly calculate the quantum capacity of a class of Gaussian channels, which includes the important case of attenuation channels modeling optical fibers with photon losses and broadband channels where losses and photon number constraints might be frequency dependent. Along the way we provide two tools that might be of independent interest: a complete characterization of degradable Gaussian channels and of those arising from teleporting through Gaussian states.

Preliminaries.—Before we derive the main results we will briefly recall the basic notions [6,7]. Consider a bosonic system of N modes characterized by N pairs of canonical operators $(Q_1, P_1, \ldots, Q_N, P_N) =: R$ for which the commutation relations $[R_k, R_l] = i\sigma_{kl}$ are represented by the symplectic matrix $\sigma = \oplus_{k=1}^{N}(i\sigma_y)$. The exponentials $W_\xi := e^{i\xi R}$, $\xi \in \mathbb{R}^{2N}$ are called Weyl displacement operators. Their expectation value, the characteristic function, $\chi(\xi) := \mathrm{Tr}[\rho W_\xi]$ is the Fourier transform of the Wigner function and for Gaussian states

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

with first moments $d_k = \mathrm{Tr}[\rho R_k]$ and covariance matrix (CM) $\gamma_{kl} := \mathrm{Tr}[\rho\{R_k - d_k, R_l - d_l\}_+]$. Note that coherent, squeezed and thermal states in quantum optics are all Gaussian states.

Gaussian channels [7,8] transform Weyl operators as $W_\xi \mapsto W_{X\xi}e^{-(1/4)\xi Y\xi}$ and act on covariance matrices as

$$\gamma \mapsto X^T\gamma X + Y. \tag{2}$$

Particularly important instances of single-mode Gaussian channels are attenuation and amplification channels for which $X = \sqrt{\eta}\mathbb{1}$ and $Y = |\eta - 1|\mathbb{1}$. For $0 \le \eta \le 1$ this models a single mode of an optical fiber with transmissivity η where the environment is assumed to be in the vacuum state. The latter reflects the fact that thermal photons with optical frequencies are negligible at room temperature. For $\eta > 1$ the channel becomes an amplification channel, where the noise term Y is now a consequence of the Heisenberg uncertainty.

Teleportation channels.—We will now derive the form of Gaussian channels which are obtained when teleporting through a centered bipartite Gaussian state. As this is useful for applying but not necessary for understanding the following it might be skipped by the reader. Let

$$\Gamma = \begin{pmatrix} \Gamma_A & \Gamma_C \\ \Gamma_C^T & \Gamma_B \end{pmatrix}$$

be the CM of a Gaussian state of $N_A + N_B$ modes with $N_A = N_B$. Assume Bob wants to teleport a quantum state of N_B modes with CM γ to Alice. Using the standard protocol [9] he sends pairs of modes from γ and Γ_B through 50:50 beam splitters, measures the Q and P quadratures, and then communicates the outcomes. Depending on the latter Alice applies displacements to the modes in Γ_A. The simplest way of deriving an expression for the output is to start with the Wigner representation and to assume that the state to be teleported is a centered Gaussian. The Wigner function before the measurement is up to normalization given by $\exp - \xi[M_{\rm BS}^T(\Gamma \oplus \gamma)^{-1}M_{\rm BS}]\xi$, where $M_{\rm BS}$ corresponds to the beam-splitter operation. With $\xi = (\xi_A, \xi_B, \xi_{B'})$ the final Wigner function is then proportional to

$$\int d\xi_B d\xi_{B'} e^{-\xi[M_X^T M_{\rm BS}^T(\Gamma\oplus\gamma)^{-1}M_{\rm BS}M_X]\xi}, \tag{3}$$

where M_X incorporates the displacements, i.e., it is the identity matrix plus an arbitrary $2N_B \times 2N_B$ off-diagonal block which maps the $2N_B$ measurement outcomes onto the respective displacements. To circumvent integrating Eq. (3) we can now go to the characteristic function, i.e., the Fourier transformed picture. The integration then boils down to picking out the upper left block of the inverted matrix $[M_X^T M_{\rm BS}^T(\Gamma \oplus \gamma)^{-1}M_{\rm BS}M_X]^{-1}$. The inversion is, however, trivial since $M_{\rm BS}^{-1} = M_{\rm BS}^T$ and M_X^{-1} is obtained from M_X by changing the sign of all off-diagonal entries. In this way we obtain that the input CM is transformed to

$$\gamma \mapsto X^T\gamma X + [\Gamma_A + \Gamma_C\Lambda X + (\Gamma_C\Lambda X)^T + X^T\Lambda^T\Gamma_B\Lambda X], \tag{4}$$

where $\Lambda = \mathrm{diag}(1, -1, 1, -1, \ldots)$ and X is such that $\sqrt{2}X$ is the matrix of displacement transformations, i.e., the gain which is typically chosen to be $\sqrt{2}\mathbb{1}$.

Clearly, Eq. (4) has the form (2) and following the above lines it is straight forward to show that the channel is Gaussian and maps any (not necessarily centered Gaussian) input characteristic function $\chi_{\rm in}$ into

$$\chi_{\rm out}(\xi) = \chi_{\rm in}(X\xi)\chi_\Gamma(\xi \oplus \Lambda X\xi). \tag{5}$$

For standard protocols ($X = \mathbb{1}$) on single modes ($N_A = N_B = 1$) this was derived in [10].

Achievable rates for arbitrary channels.—The subject of interest is the quantum capacity $Q(T)$ of an arbitrary, *a priori* unknown, channel T. We will show how one can certify achievable rates for the transmission of quantum information through T by only looking at the CM Γ of a state $\rho_T = (T \otimes \mathrm{id})(\psi)$ which is obtained by sending half of an arbitrary entangled state ψ through the channel. Γ could be determined by homodyne measurements. The argument combines (i) the relation between entanglement distillation and quantum capacities observed in [11], (ii) the extremality of Gaussian states shown in [12], and (iii) the explicit form of Gaussian teleportation channels derived in the previous section. All together this leads to the chain of inequalities

$$Q(T) \geq D_{\leftarrow}(\rho_T) \geq D_{\leftarrow}(\mathcal{G}(\rho_T)) \geq Q(T_G). \tag{6}$$

Here $D_{\leftarrow}(\rho_T)$ is the distillable entanglement under protocols with one-way communication (from Bob to Alice). Since a classical side channel does not increase $Q(T)$ this is clearly a lower bound to the capacity as Alice and Bob could simply first distill ρ_T and then use the obtained maximally entangled states for teleportation [11]. The second inequality uses that replacing ρ_T by a Gaussian state $\mathcal{G}(\rho_T)$ with the same CM Γ can only decrease the distillable entanglement [12] (see Fig. 1 for an operational meaning). Finally, if we use the Gaussian state in turn as a resource for establishing a teleportation channel T_G we end up with the sought inequality $Q(T) \geq Q(T_G)$. T_G is then the Gaussian channel in Eqs. (4) and (5), which is for a fixed teleportation protocol (a fixed matrix X) completely determined by Γ.

Bounds on the quantum capacity of Gaussian channels were derived in [13,14] and we will show below that it can be calculated exactly for some important cases. Note that a simple bound for $Q(T)$ can be obtained from a lower bound to $D_{\leftarrow}[\mathcal{G}(\rho_T)]$, the conditional entropy of the Gaussian state with CM Γ, i.e., $Q(T) \geq S(\Gamma_A) - S(\Gamma)$.

Before we proceed, two comments on the quality of the above bound and its operational meaning are in order: The given argument holds for arbitrary T and ψ. However, since we bound by Gaussian quantities the inequality might

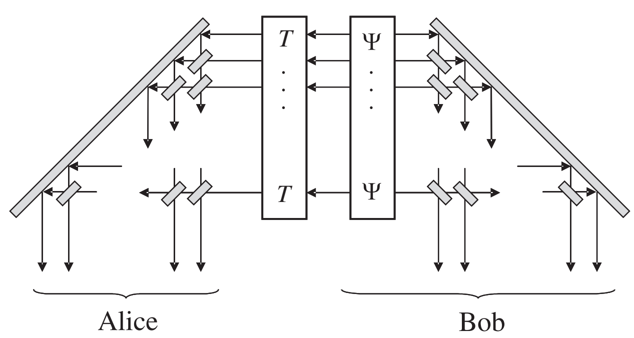

FIG. 1. In order to obtain a Gaussian channel from an arbitrary quantum channel T Bob (the sender) prepares n instances of an entangled state ψ half of which he sends through $T^{\otimes n}$. After applying two arrays of 50:50 beam splitters to the output $\rho_T^{\otimes n} = [(T \otimes \mathrm{id})(\psi)]^{\otimes n}$ the n reduced states will converge to a Gaussian state $\mathcal{G}(\rho_T)$ (with the same CM as ρ_T) which can in turn be used to establish a Gaussian teleportation channel T_G.

become trivial [i.e., $Q(T_G) = 0$ though $Q(T) \gg 0$] if both T and ψ are too far from being Gaussian. On the other hand, if T is Gaussian and $|\psi\rangle = (\cosh r)^{-1}\sum_n (\tanh r)^n |nn\rangle$ is a two-mode squeezed state, then in the limit $r \to \infty$ the inequality becomes tight, i.e., $Q(T_G) \to Q(T)$ with exponentially vanishing gap. This also indicates how (bounds on) the rate achievable by a given channel can be probed experimentally: sending part of a two-mode squeezed state through T and measuring the second moments of the resulting state allows to compute $Q(T_G)$ using the formulas below.

Quantum capacity of Gaussian channels.—It was proven in [2] that the quantum capacity of a quantum channel T can be expressed as

$$Q(T) = \lim_{n\to\infty} \frac{1}{n} \sup_\rho J(\rho, T^{\otimes n}), \tag{7}$$

$$J(\rho, T) = S(T(\rho)) - S((T \otimes \mathrm{id})(\psi)), \tag{8}$$

where ψ is a purification of ρ and J is known as the coherent information. In general, the calculation of $Q(T)$ from the above formula is a daunting task since (i) the coherent information is known to be not additive, i.e., the regularization $n \to \infty$ is necessary, and (ii) due to lacking concavity properties there are local maxima which are not global ones. On top of this, for bosonic channels the optimization is over an infinite dimensional space.

Fortunately, for a class of Gaussian channels including the important case of the lossy channel, these obstacles can be circumvented by exploiting recent results on degradability of channels [3,15] and extremality of Gaussian states [12].

To this end consider a channel $T(\rho_S) = \mathrm{Tr}_E[U(\rho_S \otimes \varphi_E)U^\dagger]$ expressed in terms of a unitary coupling between the system S and the environment E which is initially in a pure state φ_E. The conjugate channel $T_c(\rho_S) = \mathrm{Tr}_S[U(\rho_S \otimes \varphi_E)U^\dagger]$ is defined as a mapping from the system to the environment. As shown in [3] the coherent information can be expressed in terms of a conditional entropy if there exists a channel T' such that $T' \circ T = T_c$; in this case T is called degradable. More precisely, if $\tilde\rho_{S'E'}$ is the extension of the state $\tilde\rho_{S'} = T' \circ T(\rho)$ to the environment E' of T', then

$$J(\rho, T) = S(\tilde\rho_{S'E'}) - S(\tilde\rho_{S'}) =: S(E'|S'). \tag{9}$$

The conditional entropy $S(E'|S')$ is known to be strongly subadditive [1], i.e., for a composite system $S(E'_{12}|S'_{12}) \le S(E'_1|S'_1) + S(E'_2|S'_2)$. This has important consequences: for a set $\{T_i\}$ of degradable channels $J(\rho, \otimes_i T_i) \le \sum_i J(\rho_i, T_i)$, where ρ_i are the corresponding reduced states, and if each T_i is a Gaussian channel, we have in addition

$$J(\rho, \otimes_i T_i) \le \sum_i J(\rho_i, T_i) \le \sum_i J(\mathcal{G}(\rho_i), T_i). \tag{10}$$

The last inequality follows from the extremality of Gaussian states with respect to the conditional entropy [7,12] together with the fact that for Gaussian channels T_c can be chosen to be Gaussian and the CM is transformed irrespective of whether the input was Gaussian or not. As a consequence, if T_i are degradable Gaussian channels, then

$$Q(\otimes_i T_i) = \sum_i \sup_{\rho_G} J(\rho_G, T_i), \tag{11}$$

where the supremum is now taken only over Gaussian input states ρ_G. Calculating the latter for Gaussian channels is now a feasible task which was solved for the single-mode case in [13] and in [14] for broadband channels under power constraints using Lagrange multipliers. In fact, if we impose a constraint on the input energy of the form $\sum_i \omega_i N_i = \mathcal{E}$, where N_i is the average input photon number of mode i with corresponding frequency ω_i, then the above argumentation still holds, since the constraint just depends on the CM. The importance of Eq. (11) stems from the fact that a large class of Gaussian channels is indeed degradable, as shown in [15] and extended below. In particular, we can apply Eq. (11) to attenuation (amplification) channels with transmissivity η (gain $\sqrt{\eta}$). Together with the optimization carried out in [13] [Eq. (5.9)] this yields (see Fig. 2)

$$Q(\eta) = \max\{0, \log_2|\eta| - \log_2|1-\eta|\}. \tag{12}$$

Note that the quantum capacity of every degradable Gaussian channel can easily be calculated as J becomes a concave function of the CM such that local maxima are global ones.

Degradable Gaussian channels.—We will now investigate the condition under which Eq. (11) was derived and characterize the set of degradable Gaussian channels, extending the results of [15]. To this end we represent the channel in terms of a unitary coupling between the system with N_S modes and a (minimally represented) environment of $N_E \le 2N_S$ modes which are initially in the vacuum state with *CM* $\gamma_E = \mathbb{1}$. The interaction is described by a sym-

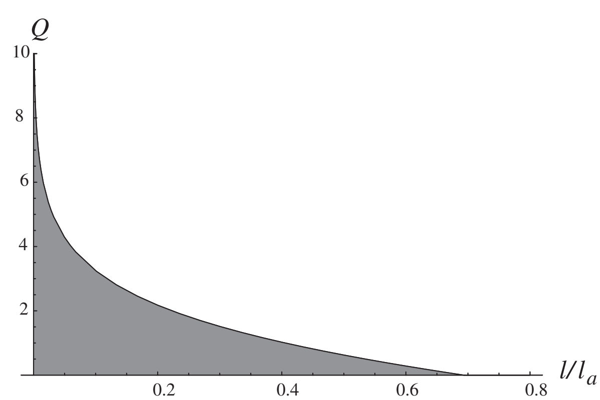

FIG. 2. Quantum capacity of a channel with photon losses as a function of the transmission length l in terms of the absorption length l_a, i.e., $\eta = e^{-l/l_a}$. For quantum memories l and l_a are storage and decay time. The capacity vanishes for $l/l_a = \ln 2 \approx 0.693$, where the channel can be considered to be part of a symmetric approximate cloning channel.

plectic matrix of size $2(N_E + N_S) \times 2(N_E + N_S)$ which we write in block form as

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

[16]. The output CM of the channel T: $\gamma \mapsto X\gamma X^T + Y$ is then simply the lower right block of $S(\gamma_E \oplus \gamma)S^T$ (i.e., $D = X$ and $C\gamma_E C^T = Y$), whereas the conjugate channel T_c corresponds to the upper left block.

Let us first focus on the case $N_S = N_E$ and assume for simplicity that the blocks in S are nonsingular. A channel is degradable if $T_c \circ T^{-1}$ is completely positive which is for a Gaussian trace preserving map equivalent to the condition [8]

$$Y + i\sigma \geq iX\sigma X^T. \tag{13}$$

Inserting the above block structure and using [16] shows that complete positivity of $T_c \circ T^{-1}$ is equivalent to

$$0 \leq (\mathbb{1} + i\sigma) - K(\mathbb{1} + i\sigma)K^T, \qquad K = C^T D^{-T} \sigma D^{-1} C. \tag{14}$$

Expressing this in terms of X and Y finally gives [16]

$$(2X\sigma X^T \sigma^T - \mathbb{1})Y \geq 0. \tag{15}$$

Similarly we can derive a condition for degradability of T_c (antidegradability of T) which is again given by the expressions (14) and (15) which have then to be negative instead of positive semidefinite.

Since for $N_E = N_S = 1$ X is a 2×2 matrix and thus $X\sigma X^T \sigma^T = \mathbb{1} \det X$, condition (15) implies that either T or T_c is degradable, as shown in [15]. Hence, as antidegradable channels have zero quantum capacity (due to the no-cloning theorem), the quantum capacity of every Gaussian channel with $N_S = N_E = 1$ can easily be calculated. In fact, by using the freedom of acting unitarily before and after the channel (which does not change its capacity) one can generically bring the channel to a normal form [17,18] which only depends on the symplectic invariant $\det X$ such that $Q(T)$ of every such channel is given by Eq. (12) with $\eta = \det X$.

Let us finally briefly comment on the case $N_E \neq N_S$. If the environment is smaller than the system, then we can easily follow the above lines for instance by choosing a representation of the channel with larger N_E equal to N_S [19]. It is worth mentioning that if S corresponds to a passive (i.e., number preserving) operation, then for $N_E < N_S$ there are always unaffected modes such that $Q(T) = \infty$ without additional constraints. If $N_E > N_S$ then Eq. (15) is merely a necessary, whereas Eq. (14) is still a necessary and sufficient condition for degradability [19]. Applying the latter to a general single-mode channel with $N_S = 1$, $N_E = 2$ shows that generically one has neither degradability nor antidegradability. Hence, it remains open whether in this case the capacity is given by Eq. (11). However, we can easily derive an upper bound by exploiting the fact that every Gaussian channel T can be decomposed as $T = T_1 \circ T_2$, where T_2 is a minimal noise channel [8] for which $N_E = N_S$ with $X_2 = X$, $Y_2 \leq Y$ and T_1 is a classical noise channel for which $X_1 = \mathbb{1}$, $Y_1 = Y - Y_2$. Because of the bottleneck inequality for capacities (cf. [13]) we have $Q(T) \leq Q(T_2)$ where the latter is in the single-mode case again given by Eq. (12) with $\eta = \det X$. A lower bound is always given by the right-hand side of Eq. (11) as calculated in [13].

We appreciate discussions with J. I. Cirac and M. E. Shirokov and financial support by Spanish projects Ramon y Cajal and MTM2005-00082.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).
[2] P. W. Shor, in The Quantum Channel Capacity and Coherent Information: Lecture Notes, MSRI Workshop on Quantum Computation, 2002 (unpublished); I. Devetak, IEEE Trans. Inf. Theory **51**, 44 (2005); S. Lloyd, Phys. Rev. A **55**, 1613 (1997).
[3] I. Devetak and P. W. Shor, quant-ph/0311131.
[4] A. Wallraff *et al.*, Nature (London) **431**, 162 (2004); A. Blais *et al.*, Phys. Rev. A **69**, 062320 (2004).
[5] D. F. Phillips *et al.*, Phys. Rev. Lett. **86**, 783 (2001); C. Liu *et al.*, Nature (London) **409**, 490 (2001); B. Julsgaard *et al.*, Nature (London) **432**, 482 (2004); T. Chanelière *et al.*, Nature (London) **438**, 833 (2005).
[6] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland Publishing, Amsterdam, 1982).
[7] J. Eisert and M. M. Wolf, quant-ph/0505151.
[8] G. Lindblad, J. Phys. A **33**, 5059 (2000); B. Demoen, P. Vanheuverzwijn, and A. Verbeure, Lett. Math. Phys. **2**, 161 (1977).
[9] L. Vaidman, Phys. Rev. A **49**, 1473 (1994); S. L. Braunstein and H. J. Kimble, Phys. Rev. Lett. **80**, 869 (1998).
[10] P. Marian, T. A. Marian, and H. Scutaru, Rom. J. Phys. **48**, 727 (2003).
[11] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[12] M. M. Wolf, G. Giedke, and J. I. Cirac, Phys. Rev. Lett. **96**, 080502 (2006).
[13] A. S. Holevo and R. F. Werner, Phys. Rev. A **63**, 032312 (2001).
[14] V. Giovannetti, S. Lloyd, L. Maccone, and P. W. Shor, Phys. Rev. A **68**, 062323 (2003).
[15] F. Caruso and V. Giovannetti, Phys. Rev. A **74**, 062307 (2006).
[16] See footnote in Michael M. Wolf, David Perez-Garcia, Geza Giedke, quant-ph/0606132.
[17] A. S. Holevo, quant-ph/0607051.
[18] A. Serafini, J. Eisert, and M. M. Wolf, Phys. Rev. A **71**, 012320 (2005).
[19] The only assumption made in the derivation of condition (14) is nonsingularity of the diagonal blocks A and D.