Evaluating capacities of bosonic Gaussian channels

A. S. Holevo* and R. F. Werner†
Institut für Mathematische Physik, TU Braunschweig, Mendelssohnstr. 3, 38106 Braunschweig, Germany

(Received 15 December 1999; published 15 February 2001)

We show how to compute or at least to estimate various capacity-related quantities for bosonic Gaussian channels. Among these are the coherent information, the entanglement-assisted classical capacity, the one-shot classical capacity, and a quantity involving the transpose operation, shown to be a general upper bound on the quantum capacity, even allowing for finite errors. All bounds are explicitly evaluated for the case of a one-mode channel with attenuation or amplification and classical noise.

DOI: 10.1103/PhysRevA.63.032312 PACS number(s): 03.67.Hk, 03.65.Db, 42.50.−p

I. INTRODUCTION

In the past several years impressive progress was achieved in the understanding of the classical and quantum capacities of quantum communication channels (see, in particular, Refs. [1–6], where the reader can also find further references). It appears that a quantum channel is characterized by a whole variety of different capacities depending both on the kind of the information transmitted and the specific protocol used.

Most of this literature studies the properties of systems and channels described in finite dimensional Hilbert spaces. Recently, however, there has been a burst of interest (see e.g., Ref. [7]) in other kinds of systems, sometimes called the ‘‘continuous variable’’ quantum systems, whose basic variables satisfy Heisenberg’s canonical commutation relations (CCR). There are two reasons for this new interest. On the one hand, such systems play a central role in quantum optics, the canonical variables being the quadratures of the field. Therefore some of the current experimental realizations [8] of quantum information processing are carried out in such systems. In particular, the bosonic Gaussian channels studied in this paper can be seen as basic building block of quantum optical communication systems, allowing us to build up complex operations from ‘‘easy, linear’’ ones and a few basic ‘‘expensive, nonlinear’’ operations, such as squeezers and parametric down converters.

The other reason for the interest in these systems is that in spite of the infinite dimension of their underlying Hilbert spaces they can be handled with techniques from finite-dimensional linear algebra, much in analogy to the finite-dimensional quantum systems on which the pioneering work on quantum information was done. Roughly speaking this analogy replaces the density matrix by the covariance matrix of a Gaussian state. Then operations like the diagonalization of density matrices, the Schmidt decomposition of pure states on composite systems, the purification of mixed states, the computation of entropies, and the partial transpose operation on states and channels, which are familiar from the usual finite-dimensional setup, can be expressed once again by operations on finite-dimensional matrices in the continuous variable case. The basic framework for doing all this is not new, and goes under the heading ‘‘phase-space quantum mechanics’’ or, in the quantum field theory and statistical mechanics communities, ‘‘quasifree Bose systems’’ [9]. Both authors of this paper have participated in the development of this subject a long time ago [10–12]. In this paper, continuing [13] and [14], we make further contributions to the study of information properties of linear bosonic Gaussian channels. We focus on the aspects essential for physical computations and leave aside a number of analytical subtleties related to infinite dimensionality and unboundedness unavoidably arising in connection with bosonic systems and Gaussian states.

The paper is organized as follows. In Sec. II we recapitulate some notions of capacity, which are currently under investigation in the literature, and what is known about them. Naturally this cannot be a full review, but will be limited to those quantities that we will evaluate or estimate in the subsequent sections. An addition to the spectrum of capacitylike quantities is discussed in Sec. II B: an upper bound on the quantum capacity (even allowing finite errors), which is both simple to evaluate and remarkably close to maximized coherent information, a bound conjectured to be exact. In Sec. III we summarize the basic properties of Gaussian states. Although our main topic is channels, we need this to get an explicit handle on the purification operation, which is needed to compute the entropy exchange, and hence all entropy based capacities. Bosonic Gaussian channels are studied in Sec. IV. Here we introduce the techniques for determining the capacity quantities introduced in Sec. I, deriving general formulas where possible. In Sec. V we apply these techniques to the case of a single-mode channel comprising attenuation or amplification and a classical noise. Some technical points are treated in the Appendixes.

II. NOTIONS OF CAPACITY

A. Basic entropy and information quantities

Consider a general quantum system in a Hilbert space $\mathcal{H} = \mathcal{H}_Q$. Its states are given by density operators ρ on $\mathcal{H}$. A *channel* is a transformation $\rho \rightarrow T[\rho]$ of quantum states of the system, which is given by a completely positive, trace-preserving map on trace class operators. This view of channels corresponds to the Schrödinger picture. The Heisenberg

*Permanent address: Steklov Mathematical Institute, Gubkina 8, 117966 Moscow, Russia. Electronic address: holevo@mi.ras.ru
†Electronic address: r.werner@tu-bs.de

picture is given by the dual linear operator $X \to T^*[X]$ on the observables X, which is defined by the relation

$$\mathrm{Tr}\, T[\rho] X = \mathrm{Tr}\, \rho T^*[X],$$

and has to be completely positive and unit preserving (cf. [15]).

It can be shown (see, e.g., [16]) that any channel in this sense arises from a unitary interaction U of the system with an environment described by another Hilbert space $\mathcal{H}_E$ which is initially in some state ρ_E,

$$T[\rho] = \mathrm{Tr}_E\, U(\rho \otimes \rho_E) U^*,$$

where Tr_E denotes partial trace with respect to $\mathcal{H}_E$, and vice versa. The representation is not unique, and the state ρ_E can always be chosen pure, $\rho_E = |\psi_E\rangle\langle\psi_E|$. The definition of the channel has obvious generalization to the case where input and output are described by different Hilbert spaces.

Let us denote by

$$H(\rho) = -\mathrm{Tr}\, \rho \log \rho \tag{2.1}$$

the von Neumann entropy of a density operator ρ. We call ρ the input state, and $T[\rho]$ the output state of the channel. There are three important entropy quantities related to the pair (ρ, T), namely, the entropy of the input state $H(\rho)$, the entropy of the output state $H(T[\rho])$, and the entropy exchange $H(\rho, T)$. While the definition and the meaning of the first two entropies is clear, the third quantity is somewhat more sophisticated. To define it, one introduces the *reference system*, described by the Hilbert space $\mathcal{H}_R$, isomorphic to the Hilbert space $\mathcal{H}_Q = \mathcal{H}$ of the initial system. Then according to Refs. [17] and [3], there exists a *purification* of the state ρ, i.e., a unit vector $|\psi\rangle \in \mathcal{H}_Q \otimes \mathcal{H}_R$ such that

$$\rho = \mathrm{Tr}_R |\psi\rangle\langle\psi|.$$

The *entropy exchange* is then defined as

$$H(\rho, T) = H((T \otimes \mathrm{id})[|\psi\rangle\langle\psi|]), \tag{2.2}$$

that is, as the entropy of the output state of the dilated channel $(T \otimes \mathrm{id})$ applied to the input which is purification of the state ρ. Alternatively,

$$H(\rho, T) = H(\rho'_E),$$

where $\rho'_E = T_E[\rho]$ is the final state of the environment, and the channel T_E from $\mathcal{H}_Q$ to $\mathcal{H}_E$ is defined as

$$T_E[\rho] = \mathrm{Tr}_Q\, U(\rho \otimes \rho_E) U^*,$$

provided the initial state ρ_E of the environment is pure [17,3].

From these three entropies one can construct several information quantities. In analogy with classical information theory, one can define *quantum mutual information* between the reference system R (which mirrors the input Q) and the output of the system Q' [17,4] as

$$\begin{aligned} I(\rho, T) &= H(\rho'_R) + H(\rho'_Q) - H(\rho'_{RQ}) \\ &= H(\rho) + H(T[\rho]) - H(\rho, T). \end{aligned} \tag{2.3}$$

The quantity $I(\rho, T)$ has a number of ''natural'' properties, in particular, positivity, concavity with respect to the input state ρ and additivity for parallel channels [4]. Moreover, the maximum of $I(\rho, T)$ with respect to ρ was argued recently to be equal to the *entanglement-assisted classical capacity* of the channel [6,18], namely, the classical capacity of the superdense coding protocol using the noisy channel T. It was shown that this maximum is additive for parallel channels, the one-shot expression thus giving the full (asymptotic) capacity.

It would be natural to compare this quantity with the (unassisted) classical capacity $C(T)$ (the definition of which is outlined in Sec. II B); however, it is still not known whether this capacity is additive for parallel channels. This makes us focus on the one-shot expression, emerging from the coding theorem for classical-quantum channels [2]

$$C_1(T) = \max \left[H\left(\sum_i p_i T[\rho_i] \right) - \sum_i p_i H(T[\rho_i]) \right], \tag{2.4}$$

where the maximum is taken over all probability distributions $\{p_i\}$ and collections of density operators $\{\rho_i\}$ (possibly satisfying some additional input constraints). $C_1(T)$ is equal to the capacity of T for classical information, if the coding is required to avoid entanglement between successive inputs to the channel. The full capacity is then attained as the length n of the blocks, over which encoding may be entangled, goes to infinity, i.e.,

$$C(T) = \lim_{n \to \infty} \frac{1}{n} C_1(T^{\otimes n}). \tag{2.5}$$

An important component of $I(\rho, T)$ is the *coherent information*

$$J(\rho, T) = H(T[\rho]) - H(\rho, T), \tag{2.6}$$

the maximum of which has been conjectured to be the (one-shot) quantum capacity of the channel T [19,3]. Its properties are not so nice. It can be negative, its convexity properties with respect to ρ are not known, and its maximum was shown to be strictly superadditive for certain parallel channels [20], hence the conjectured full quantum capacity may be greater than the one-shot expression, in contrast to the case of the entanglement-assisted classical capacity. In this paper we shall also compare this expression with a new upper bound on the quantum capacity $Q(T)$ (as introduced, e.g., in Sec. II B).

B. A general bound on quantum channel capacity

In this section we will establish a general estimate on the quantum channel capacity, which will then be evaluated in the Gaussian case, and will be compared with the estimates of coherent information. Let us recall first a definition of the capacity $Q(T)$ of a general channel T for quantum information. Intuitively, it is the number of qubits that can be faithfully transmitted per use of the channel with the best possible error correction. The standard of comparison is the ideal 1-qubit channel id_2, where id_n denotes the identity map on the $n\times n$ matrices. Then the *quantum capacity* $Q(T)$ of a channel T (possibly between systems of different types) is defined as the supremum of all numbers c, which are ''attainable rates'' in the following sense: *For any pair of sequences* n_α, m_α *with* $\lim_\alpha(n_\alpha/m_\alpha)=c$ *we can find encoding operations* E_α *and decoding operations* D_α *such that*

$$\|\mathrm{id}_2^{\otimes n_\alpha}-D_\alpha T^{\otimes m_\alpha}E_\alpha\|_{\mathrm{cb}}\to 0.$$

Here $\|\cdot\|_{\mathrm{cb}}$ is the so-called ''norm of complete boundedness'' (cb) [21], which is defined as the supremum with respect to n of the norms $\|(T\otimes\mathrm{id}_n)\|$. It is equal to the ''diamond metric'' introduced in [22]. We use this norm because, on the one hand, it leads to the same capacity as analogous definitions based on other error criteria (e.g., fidelities [3,5]) and, on the other hand, it has the best properties with respect to tensor products, which are our main concern. In particular, $\|T\otimes S\|_{\mathrm{cb}}=\|T\|_{\mathrm{cb}}\cdot\|S\|_{\mathrm{cb}}$. Completely positive maps satisfy $\|T\|_{\mathrm{cb}}=\|F\|$, where $F=T^*[I]$ is the normalization operator (we denote I the unit operator). In particular, $\|T\|_{\mathrm{cb}}=1$ for any channel. We also note another kind of capacity, in which a much weaker requirement is made on the errors, namely,

$$\|\mathrm{id}_2^{\otimes n_\alpha}-D_\alpha T^{\otimes m_\alpha}E_\alpha\|_{\mathrm{cb}}\leqslant\varepsilon<1 \tag{2.7}$$

for all sufficiently large α, and some fixed ε. We call the resulting capacity the ε*-quantum capacity*, and denote it by $Q_\varepsilon(T)$. Of course, $Q(T)\leqslant Q_\varepsilon(T)$, and by analogy with the classical case (strong converse of Shannon's coding theorem) one would conjecture that equality always holds.

The unassisted classical capacity $C(T)$ can be defined similarly with the sole difference that both the domain of encodings E and the range of decodings D should be restricted to the state space of the Abelian subalgebra of operators diagonalizable in a fixed orthonormal basis. In that case there is no need to use the cb norm, as it coincides with the usual norm. According to recently proven strong converse to the quantum coding theorem [23,24], $C_\varepsilon(T)=C(T)$ where $C_\varepsilon(T)$ is defined similarly to $Q_\varepsilon(T)$.

The criterion we will formulate makes essential use of the transpose operation, which we will denote by the same letter Θ in any system. For matrix algebras, Θ can be taken as the usual transpose operation. However, it makes no difference to our considerations, if any other antiunitarily implemented symmetry (e.g., time reversal) is chosen. In an abstract C^* algebra setting Θ is best taken as the ''op'' operation, which maps every algebra to its ''opposite.'' This algebra has the same underlying vector space, but all products AB are replaced by their opposite BA. Obviously, a commutative algebra is the same as its opposite, so on classical systems Θ is the identity. Although the transpose maps density operators to density operators, it is not an admissible quantum channel, because positivity is lost, when coupling the operation with the identity transformation on other systems, i.e., Θ is not *completely* positive. A similar phenomenon happens for the norm of Θ: we have $\|\Theta\|_{\mathrm{cb}}>1$ unless the system is classical. In fact,

$$\|\Theta_n\|_{\mathrm{cb}}=n, \tag{2.8}$$

where Θ_n denotes the transposition on the $n\times n$ matrices [21]. We note that since we do not distinguish the transpose on different systems in our notation, the observation that tensor products can be transposed factor by factor is expressed by the equation $\Theta=\Theta\otimes\Theta$. Moreover, although for a channel T, the operator $T\Theta$ may fail to be completely positive, $\Theta T\Theta$ is again a channel, and, in particular, satisfies $\|\Theta T\Theta\|_{\mathrm{cb}}=1$.

The main result of this section is the estimate

$$Q_\varepsilon(T)\leqslant\log\|T\Theta\|_{\mathrm{cb}}\equiv Q_\Theta(T), \tag{2.9}$$

for any channel T. The proof is quite simple. Suppose $n_\alpha/m_\alpha\to c\leqslant Q_\varepsilon(T)$, and encoding E_α and decoding D_α are as in the definition of $Q_\varepsilon(T)$. Then by Eq. (2.8) we have

$$\begin{aligned}2^{n_\alpha}&=\|\mathrm{id}_2^{\otimes n_\alpha}\Theta\|_{\mathrm{cb}}\\&\leqslant\|(\mathrm{id}_2^{\otimes n_\alpha}-D_\alpha T^{\otimes m_\alpha}E_\alpha)\Theta\|_{\mathrm{cb}}+\|D_\alpha T^{\otimes m_\alpha}E_\alpha\Theta\|_{\mathrm{cb}}\\&\leqslant\|\Theta_{2^{n_\alpha}}\|_{\mathrm{cb}}\|\mathrm{id}_2^{\otimes n_\alpha}-D_\alpha T^{\otimes m_\alpha}E_\alpha\|_{\mathrm{cb}}\\&\quad+\|D_\alpha(T\Theta)^{\otimes m_\alpha}\Theta E_\alpha\Theta\|_{\mathrm{cb}}\\&\leqslant 2^{n_\alpha}\varepsilon+\|T\Theta\|_{\mathrm{cb}}^{m_\alpha},\end{aligned}$$

where at the last inequality we have used that D_α and $\Theta E_\alpha\Theta$ are channels, and that the cb norm is exactly tensor multiplicative, so $\|X^{\otimes m}\|_{\mathrm{cb}}=\|X\|_{\mathrm{cb}}^m$. Hence, by taking the logarithm and dividing by m_α, we get

$$\frac{n_\alpha}{m_\alpha}\log 2+\frac{\log(1-\varepsilon)}{m_\alpha}\leqslant\log\|T\Theta\|_{\mathrm{cb}}.$$

If we take base 2 logarithms, as is customary in information theory, we have $\log 2=1$. Then in the last inequality we can go to the limit $\alpha\to\infty$, obtaining $c\leqslant Q_\Theta(T)$, and Eq. (2.9) follows by taking the supremum over all attainable rates c. Note that base 2 logarithms are built into the above definition of capacity, because we are using the ideal qubit channel as the standard of comparison. This amounts only to a change of units. If another base for logarithms is chosen, this should also be done consistently in all entropy expressions, and Eq. (2.9) holds once again without additional constants.

The upper bound $Q_\Theta(T)$ computed in this way has some remarkable properties (proved in Appendix A), which make it a capacitylike quantity in its own right. For example, it is exactly additive,

$$Q_\Theta(S\otimes T)=Q_\Theta(S)+Q_\Theta(T), \tag{2.10}$$

for any pair S,T of channels, and satisfies the ''bottleneck inequality'' $Q_\Theta(ST)\leqslant\min\{Q_\Theta(S),Q_\Theta(T)\}$. Moreover, it coincides with the quantum capacity on ideal channels: $Q_\Theta(\mathrm{id}_n)=Q(\mathrm{id}_n)=\log n$, and it vanishes whenever $T\Theta$ is completely positive. In particular, $Q_\Theta(T)=0$, whenever T is *separable* in the sense that it can be decomposed as $T=PM$ into a measurement M and a subsequent preparation

P based on the measurement results. We note that Q_Θ is also closely related to the entanglement quantity $\log\|(\mathrm{id}\otimes\Theta)[\rho]\|_1$, i.e., the logarithm of the trace norm of the partial transpose of the density operator, which enjoys analogous properties.

III. QUANTUM GAUSSIAN STATES

A. Canonical variables and Gaussian states

In this section we recapitulate some results from [10,13,14] for the convenience of the reader. Our approach to quantum Gaussian states is based on the characteristic function of the state which closely parallels classical probability [11,12], and is perhaps the simplest and most transparent analytically. An alternative approach can be based on the Wigner ‘‘distribution function’’ [25].

Let q_j,p_j be the canonical observables satisfying the Heisenberg CCR

$$[q_j,p_k]=i\delta_{jk}\hbar I,\quad [q_j,q_k]=0,\quad [p_j,p_k]=0.$$

We introduce the column vector of operators

$$R=[q_1,p_1,\ldots,q_s,p_s]^T,$$

the real column $2s$ vector $z=[x_1,y_1,\ldots,x_s,y_s]^T$, and the unitary operators in $\mathcal{H}$

$$\begin{aligned}V(z)&=\exp i\sum_{j=1}^{s}(x_jq_j+y_jp_j)\\&=\exp iR^Tz.\end{aligned}\tag{3.1}$$

These ‘‘Weyl operators’’ satisfy the Weyl-Segal CCR

$$V(z)V(z')=\exp\left[\frac{i}{2}\Delta(z,z')\right]V(z+z'),\tag{3.2}$$

where

$$\Delta(z,z')=\hbar\sum_{j=1}^{s}(x_j'y_j-x_jy_j')\tag{3.3}$$

is the canonical symplectic form. The space Z of real $2s$ vectors equipped with the form $\Delta(z,z')$ is what one calls a *symplectic vector space*. We denote by

$$\Delta=\begin{bmatrix}0&\hbar&&&\\-\hbar&0&&&\\&&\ddots&&\\&&&0&\hbar\\&&&-\hbar&0\end{bmatrix}\tag{3.4}$$

the $(2s)\times(2s)$-skew-symmetric *commutation matrix* of components of the vector R, so that

$$\Delta(z,z')=-z^T\Delta z'.$$

Most of the results below are valid for the case where the commutation matrix is an arbitrary (nondegenerate) skew-symmetric matrix, not necessarily of the canonical form (3.4).

A density operator ρ has *finite second moments* if $\mathrm{Tr}(\rho q_j^2)<\infty$ and $\mathrm{Tr}(\rho p_j^2)<\infty$ for all j. In this case one can define the vector *mean* and the *correlation matrix* α by the formulas

$$m=\mathrm{Tr}\,\rho R;\quad \alpha-\frac{i}{2}\Delta=\mathrm{Tr}(R-m)\rho(R-m)^T.\tag{3.5}$$

The mean can be an arbitrary real vector. The correlation matrix α is real and symmetric. A given α is the correlation matrix of some state if and only if it satisfies the *matrix uncertainty relation*

$$\alpha-\frac{i}{2}\Delta\geqslant 0.\tag{3.6}$$

We denote by $\Sigma(m,\alpha)$ the set of states with fixed mean m and the correlation function α. The density operator ρ is called *Gaussian*, if its *quantum characteristic function* $\phi(z)=\mathrm{Tr}\,\rho V(z)$ has the form

$$\phi(z)=\exp(im^Tz-\tfrac{1}{2}z^T\alpha z),\tag{3.7}$$

where m is a column $(2s)$ vector and α is a real symmetric $(2s)\times(2s)$ matrix. One then can show that m is indeed the mean, and α is the correlation matrix, and Eq. (3.7) defines the unique Gaussian state in $\Sigma(m,\alpha)$. In what follows we will be interested mainly in the case $m=0$.

The correlation matrix α describes a quadratic form rather than an operator. Therefore its eigenvalues have no intrinsic significance, and depend on the choice of basis in Z. On the other hand, the operator $\hat{\alpha}$ defined by $z^T\alpha z=\Delta(z,\hat{\alpha}z)$ has a basis free meaning. In matrix notation it is $\hat{\alpha}=\Delta^{-1}\alpha$. This operator is always diagonalizable, and its eigenvalues come in pairs $\pm i\gamma_j$. Diagonalizing this operator is essentially the same as the *normal mode decomposition* of the phase space, when the form $z^T\alpha z$ is considered as the Hamiltonian function of a system of oscillators. It leads to a decomposition of the phase space into two-dimensional subspaces, such that on the j^{th} subspace we have (in some new canonical variables $\tilde{q}_j,\tilde{p}_j$)

$$\alpha=\hbar\begin{bmatrix}\gamma_j&0\\0&\gamma_j\end{bmatrix},\quad \Delta=\hbar\begin{bmatrix}0&1\\-1&0\end{bmatrix},\tag{3.8}$$

and all terms between different blocks vanish. The matrix uncertainty relation now requires $\gamma_j\geqslant 1/2$, in which equality holds if and only if ρ_j is the pure (minimum-uncertainty) state. Hence a general Gaussian state ρ is pure if and only if all $\gamma_j=1/2$, or

$$(\Delta^{-1}\alpha)^2=-\frac{1}{4}I,\tag{3.9}$$

in which case $\Sigma(m,\alpha)$ reduces to a single point.

B. Gauge-invariant states

We shall be interested in the particular subclass of Gaussian states most familiar in quantum optics, namely, the states having a P_- representation

$$\rho=\int |\zeta\rangle\langle\zeta|\mu_N(d^{2s}\zeta) \tag{3.10}$$

where $\mu_N(d^{2s}\zeta)$ is the complex Gaussian probability measure with zero mean and the correlation matrix N (see, e.g., [26], Sec. V.5.II). Here $\zeta\in\mathbf{C}^s$, $|\zeta\rangle$ are the coherent vectors in $\mathcal{H}$, $a|\zeta\rangle=\zeta|\zeta\rangle$, N is the positive Hermitian matrix such that

$$N=\mathrm{Tr}(a\,\rho\,a^\dagger) \tag{3.11}$$

(we use here vector notations, where $a=[a_1,\ldots,a_s]^T$ is a column vector and $a^\dagger=[a_1^\dagger,\ldots,a_s^\dagger]$ is a row vector), and $a_j=(1/\sqrt{2\hbar})(q_j+ip_j)$.

These states respect the natural complex structure in the sense that they are invariant under the gauge transformations $a\rightarrow a\exp(i\varphi)$. As shown in [13], the quantum correlation matrix of such states is

$$\alpha=\hbar\begin{bmatrix}\mathrm{Re}\,N+I/2 & -\mathrm{Im}\,N\\ \mathrm{Im}\,N & \mathrm{Re}\,N+I/2\end{bmatrix}.$$

With Pauli matrices I_2,σ_y, the real $2s\times 2s$ matrices of such form can be rewritten as complex $s\times s$ matrices, by using the correspondence

$$\begin{bmatrix}A & -B\\ B & A\end{bmatrix}=I_2A-i\sigma_yB\leftrightarrow A+iB,$$

which is an algebraic isomorphism. Obviously,

$$\frac{1}{2}\,\mathrm{Sp}\begin{bmatrix}A & -B\\ B & A\end{bmatrix}=\mathrm{Sp}(A+iB),$$

where by ‘‘Sp’’ we denote the trace of matrices, as opposed to the trace of Hilbert space operators, which is denoted by ‘‘Tr.’’ By using this correspondence, we have

$$\alpha\leftrightarrow\hbar(N+I/2),\quad \Delta\leftrightarrow-i\hbar I, \tag{3.12}$$

and

$$\Delta^{-1}\alpha\leftrightarrow i(N+I/2). \tag{3.13}$$

For the case of one degree of freedom we shall be interested in the last section, N is just a non-negative number, and ρ is an *elementary* Gaussian state with the characteristic function

$$\phi(z)=\exp\left[-\frac{\hbar}{2}\left(N+\frac{1}{2}\right)|z|^2\right], \tag{3.14}$$

where we set $|z|^2=(x^2+y^2)$. This state has a correlation matrix of the form (3.8) in the initial variables q,p, with $\gamma=N+1/2$, and is just the temperature state of the harmonic oscillator

$$\rho_\gamma=\frac{1}{\gamma+1/2}\sum_{n=0}^{\infty}\left(\frac{\gamma-1/2}{\gamma+1/2}\right)^n|n\rangle\langle n| \tag{3.15}$$

in the number basis $|n\rangle$, with the mean photon number N.

C. Computation of entropy

To compute the von Neumann entropy of a general Gaussian state one can use the normal mode decomposition. For a single mode, the density operator ρ_j with the correlation matrix (3.8), setting $\gamma_j\equiv\gamma$ for convenience, is unitarily equivalent to the state (3.15). From this one readily gets the von Neumann entropy $H(\rho_\gamma)$ by a summation of the geometric series, and for general Gaussian ρ by summing over normal modes.

To write the result in compact form, one introduces the function

$$g(x)=(x+1)\log(x+1)-x\log x,\quad x>0$$

$$g(0)=0. \tag{3.16}$$

Then

$$H(\rho)=\sum_{j=1}^{s}g\left(|\gamma_j|-\frac{1}{2}\right), \tag{3.17}$$

where γ_j runs over all eigenvalue pairs $\pm i\gamma_j$ of $\Delta^{-1}\alpha$.

One can also write this more compactly, using the following notations, which we will also use in the sequel. For any diagonalizable matrix $M=S\,\mathrm{diag}(m_j)S^{-1}$, we set $\mathrm{abs}(M)=S\,\mathrm{diag}(|m_j|)S^{-1}$, analogously for other continuous functions on the complex plane. Then Eq. (3.17) can be written as [13]

$$H(\rho)=\frac{1}{2}\,\mathrm{Sp}\,g\left(\mathrm{abs}(\Delta^{-1}\alpha)-\frac{I}{2}\right). \tag{3.18}$$

For gauge-invariant state, by using (3.13), this reduces to the well-known formula

$$H(\rho)=\mathrm{Sp}\,g(N).$$

D. Schmidt decomposition and purification

Forming a composite systems out of two systems described by CCR relations is very simple: one just joins the two sets of canonical operators, making operators belonging to different systems commute. The symplectic space of the composite system is a direct sum $Z_{12}=Z_1\oplus Z_2$, which means that elements of this space are pairs (z_1,z_2) with components $z_i\in Z_i$. In terms of Weyl operators one can write $V_{12}(z_1,z_2)=V_1(z_1)\otimes V_2(z_2)$. By definition, the symplectic matrix Δ_{12} is block diagonal with respect to the decomposition $Z=Z_1\oplus Z_2$. However, the correlation matrix α_{12} is block diagonal if and only if the state is a product. The restriction of a bipartite Gaussian state ρ to the first factor is determined by the expectations of the Weyl operators $V_1(z_1)\otimes\mathbf{1}=V_{12}(z_1,0)$, hence according to (3.7), by the correlation matrix α_1 with $z_1^T\alpha_1 z_1=(z_1,0)^T\alpha_{12}(z_1,0)$, which is just the first diagonal block in the block matrix decomposition

$$\alpha_{12}=\begin{bmatrix}\alpha_1 & \beta\\ \beta^T & \alpha_2\end{bmatrix},\quad \Delta_{12}=\begin{bmatrix}\Delta_1 & 0\\ 0 & \Delta_2\end{bmatrix}. \tag{3.19}$$

As in the case of bipartite systems with finite-dimensional Hilbert spaces there is a canonical form for *pure* states of the composite system, the Schmidt decomposition. Like the diagonalization of a one-site density operator, it can be carried out for Gaussian states at the level of correlation matrices. By writing out equation (3.9) in block matrix form, we find in particular that

$$(\Delta_1^{-1}\alpha_1)(\Delta_1^{-1}\beta)=-(\Delta_1^{-1}\beta)(\Delta_2^{-1}\alpha_2). \tag{3.20}$$

Thus $(\Delta_1^{-1}\beta)$ maps eigenvectors of $(\Delta_2^{-1}\alpha_2)$ into eigenvectors of $(\Delta_1^{-1}\alpha_1)$, with the opposite eigenvalue. Hence the spectra of the restrictions are synchronized much in the same way as in the finite-dimensional case, and all the matrices α_1,α_2,β can be diagonalized simultaneously by a suitable choice of canonical coordinates. Evaluating also the diagonal part of Eq. (3.9), one gets an equation for β, so that finally α_{12} is decomposed into blocks corresponding to (a) pure components belonging to only one subsystem, and not correlated with the other, and (b) blocks of a standard form, which can be written like Eq. (3.19) with $\alpha_1=\alpha_2=\alpha$, $\Delta_1=\Delta_2=\Delta$ from Eq. (3.8), and

$$\beta=\hbar\sqrt{\gamma^2-\frac{1}{4}}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}. \tag{3.21}$$

The purification of a general Gaussian state can easily be read off from this, by constructing such a standard form for every normal mode. In order to write β in operator form without explicit reference to the normal mode decomposition, it is most convenient to perform an appropriate reflection in the space Z_2, by which β becomes purely off-diagonal. Then we can choose [27] $\Delta_1=\Delta=-\Delta_2$ and $\alpha_2=\alpha_1=\alpha$, resulting in

$$\beta=-\beta^T=\Delta\sqrt{-(\Delta^{-1}\alpha)^2-I/4}. \tag{3.22}$$

This also covers cases with $\beta=0$ for some modes where, strictly speaking, no purification would have been necessary. We thus have

$$\Delta_{12}^{-1}\alpha_{12}=\begin{bmatrix}\Delta^{-1}\alpha & \sqrt{-(\Delta^{-1}\alpha)^2-I/4}\\ \sqrt{-(\Delta^{-1}\alpha)^2-I/4} & -\Delta^{-1}\alpha\end{bmatrix}. \tag{3.23}$$

In the gauge-invariant case, we can use the correspondence

$$\Delta_{12}^{-1}\alpha_{12}\leftrightarrow\begin{bmatrix}i(N+I/2) & \sqrt{N^2+N}\\ \sqrt{N^2+N} & -i(N+I/2)\end{bmatrix}, \tag{3.24}$$

following from Eq. (3.12).

IV. LINEAR BOSONIC CHANNELS

A. Basic properties

The characteristic property of the channels considered in this paper is their simple description in terms of phase-space structures. The key feature is that Weyl operators go into Weyl operators, up to a factor. That is, the channel map in the Heisenberg picture is of the form

$$T^*(V'(z'))=V(K^Tz')f(z'), \tag{4.1}$$

where $K:Z\rightarrow Z'$ is a linear map between phase spaces with symplectic forms Δ and Δ', respectively, and $f(z')$ is a scalar factor satisfying certain positive definiteness condition to be discussed later. Because of the linearity of K, such channels are called *linear bosonic channels* [15], and if, in addition, the factor f is Gaussian, T will be called a *Gaussian channel*. In terms of characteristic functions, Eq. (4.1) can be written as

$$\phi'(z')=\phi(K^Tz')f(z'), \tag{4.2}$$

where ϕ and ϕ' are the characteristic functions of input state ρ and output state $T[\rho]$, respectively.

We will make use of the following key properties.

(a) The dual of a linear bosonic channel transforms any polynomial in the operators R' into a polynomial in the R of the same order, provided the function f has derivatives of sufficiently high order. This property follows from the definition of moments by differentiating the relation (4.1) at the point $z'=0$.

(b) A Gaussian channel transforms Gaussian states into Gaussian states. This follows from the definition of Gaussian state and the relation (4.2).

(c) Linear bosonic channels are covariant with respect to phase-space translations. That is if $\rho^z=V(-\Delta^{-1}z)\rho V(-\Delta^{-1}z)^*$ is a shift of ρ by z, $T[\rho]$ is similarly shifted by Kz.

There is a dramatic difference in the capacities of a Gaussian channel for classical as opposed to quantum information. Classical information can be coded by using phase-space translates of a fixed state as signal states, so the output signals will also be phase-space translates of each other. Then no matter how much noise the channel may add, if we take the spacing of the input signals sufficiently large, the output states will also be sufficiently widely spaced to be distinguishable with near certainty. Therefore the unconstrained *classical capacity is infinite*. The same would be true, of course, for a purely classical channel with Gaussian noise. The classical capacity of such channels becomes an interesting quantity, however, when the ''input power'' is taken to be constrained by a fixed value, which we must take as one of the parameters defining the channel. Then arbitrarily wide spacing of input signals is no longer an alternative, because an intrinsic scale for this spacing has been introduced.

The remarkable fact of quantum information on Gaussian channels is that such an intrinsic scale is already there: it is given by $\hbar$. As we will show, the quantum information capacity is typically bounded even without an energy constraint. Loosely speaking, although we send arbitrarily many well distinguishable quantum signals through the channel, coherence in the form of commutator relations is usually lost. Surprisingly, in spite of the infinite classical capacity, the *capacity for quantum information may be zero*, which means that even joining arbitrarily many parallel channels with poor coherence properties is not good enough for sending a single qubit. This phenomenon will be explained in some detail in Sec. V.

The choice of the scalar function $f(z')$ is crucial for the quantum transmission properties of the channel. Normalization of T requires that $f(0)=1$, and it is clear that $|f(z')|\leqslant 1$ for all z', from taking norms in (4.1). Beyond that, it is not so easy to see which choices of f are compatible with the complete positivity. If f decays rapidly, T^* maps most operators to operators near the identity, which means that there is very much noise. On the other hand, there will be a lower limit to the noise, depending on the linear transformation K. Only when K is a symplectic linear map and T is reversible, the choice $f(z)\equiv 1$ is possible. Otherwise, there is some unavoidable noise.

There are two basic approaches to the determination of the admissible functions f. The first is the familiar constructive approach already used in Sec. II, based on coupling the system to an environment, a unitary evolution and subsequent reduction to a subsystem, with all of these operations in their linear bosonic Gaussian form. Basically this reduces the problem to linear transformations of systems of canonical operators. This will be described in Sec. IV B, and used for the calculation of entropy exchange in Sec. IV C. Alternatively, one can describe the admissible functions f by a twisted positive definiteness condition, and this will be used for evaluating the bound $C_\Theta(T)$ in Sec. IV D.

B. Bosonic channels via transforming canonical operators

Let R,R_E be vectors of canonical observables in $\mathcal{H},\mathcal{H}_E$, with the commutation matrices Δ,Δ_E. Consider the linear transformation

$$R'=KR+K_ER_{E,} \tag{4.3}$$

where K,K_E are real matrices (to simplify notations we write R,R_E instead of $R\otimes I_E, I\otimes R_E$, etc.) Then the commutation matrix and the correlation with respect to R' are computed via (3.5) with $m=0$, namely,

$$\alpha'-\frac{i}{2}\Delta'=\mathrm{Tr}\,R'\rho'R'^T.$$

We apply this to the special case $\rho'=\rho\otimes\rho_E$, where ρ_E and ρ are density operators in $\mathcal{H}_E$ and $\mathcal{H}$ with the correlation matrices α_E and α, respectively. Then using (4.3), we obtain

$$\Delta'=K\Delta K^T+K_E\Delta_EK_E^T,$$

$$\alpha'=K\alpha K^T+K_E\alpha_EK_E^T. \tag{4.4}$$

Of course, the operators R' need not form a complete set of observables in $\mathcal{H}\otimes\mathcal{H}_E$, but in any case α' is the correlation matrix of a system containing just the canonical variables R', and it is this state which we will consider as the output state of the channel.

For fixed state ρ_E (state of the ‘‘environment’’) the channel transformation taking the input state ρ to the output ρ' is described most easily in terms of characteristic functions,

$$\phi'(z')=\phi(K^Tz')\phi_E(K_E^Tz'). \tag{4.5}$$

We can write this as a linear Bosonic channel in the form (4.2) with

$$f(z')=\phi_E(K_E^Tz')=\mathrm{Tr}\,\rho_EV_E(K_E^Tz'). \tag{4.6}$$

Thus the factor f is expressed in terms of the characteristic function of the initial state of the environment. Obviously, the channel is Gaussian if and only if this state is Gaussian.

If we want to get the state of the environment after the channel interaction, as required in the definition of exchange entropy, we have to supplement the linear equation (4.3) by a similar equation specifying the environment variables R_E' after the interaction

$$R'=KR+K_ER_{E,}$$

$$R_E'=LR+L_ER_{E,}.$$

Assuming that $Z=Z'$ and $\Delta'=\Delta$, one can always choose L,L_E such that the combined transformation is canonical, i.e., preserves the commutation matrix

$$\begin{bmatrix}\Delta & 0\\ 0 & \Delta_E\end{bmatrix}.$$

Then the channel $T_E:\rho\rightarrow\rho_E'$ can be defined by the relation

$$T_E^*[V_E(z_E)]=V(L^Tz_E)\cdot\phi_E(L_E^Tz_E),$$

and is thus also linear bosonic.

C. Maximization of mutual information

The estimate for the entanglement-assisted classical capacity suggested by [18] is the maximum of the quantum mutual information (2.3) over all states satisfying an appropriate energy constraint. Evaluating this maximum becomes possible by the following result:[1] *Let T be a Gaussian channel. The maximum of the mutual information $I(\rho)$ over the set of states $\Sigma(m,\alpha)$ with given first and second moments is achieved on the Gaussian state.*

Proof (sketch). By purification (if necessary), we can always assume that ρ_E is pure Gaussian. Then we can write

$$I(\rho)=H(\rho)+H(T[\rho])-H(T_E[\rho]).$$

[1] The proof of this theorem was stimulated by a question posed to one of the authors (A.H.) by P. W. Shor.

Let ρ_0 be the unique Gaussian state in $\Sigma(m,\alpha)$. For simplicity we assume here that ρ_0 is nondegenerate. The general case can be reduced to this by separating the pure component in the tensor product decomposition of ρ_0. The function $I(\rho)$ is concave and its directional derivative at the point ρ_0 is (cf. [18])

$$\nabla_X I(\rho_0)=\mathrm{Tr}\,X(\ln\rho_0+I)+\mathrm{Tr}\,T[X](\ln T[\rho_0]+I)$$
$$-\mathrm{Tr}\,T_E[X](\ln T_E[\rho_0]+I).$$

By using dual maps this can be modified to

$$\nabla_X I(\rho_0)=\mathrm{Tr}\,X\{\ln\rho_0+T^*(\ln T[\rho_0])-T_E^*(\ln T_E[\rho_0])+I\}. \tag{4.7}$$

Now by property (b) of Gaussian channels, the operators $\rho_0, T[\rho_0], T_E[\rho_0]$ are (nondegenerate) Gaussian density operators, hence their logarithms are quadratic polynomials in the corresponding canonical variables (see the Appendix in [13]). By property (a) the expression in curly brackets in (4.7) is again a quadratic polynomial in R, which is a linear combination of the constraint operators in $\Sigma(m,\alpha)$. Therefore, the sufficient condition (B3) in Appendix B is fulfilled and $I(\rho)$ achieves its maximum at the point $\rho_0\in\Sigma(m,\alpha)$.

This theorem implies that the maximum of $I(\rho)$ over a set of density operators defined by arbitrary constraints on the first and second moments is also achieved on a Gaussian density operator. In particular, for an arbitrary quadratic Hamiltonian H the maximum of $I(\rho)$ over states with constrained mean energy $\mathrm{Tr}\,\rho H$ is achieved on a Gaussian state. The energy constraint is linear in terms of the correlation matrix:

$$\mathrm{Sp}(\epsilon\alpha)\leqslant N,$$

where ϵ is the diagonal energy matrix (see [13]).

When ρ and T are Gaussian, the quantities $H(\rho)$, $H(T[\rho]), H(\rho,T)$ and $I(\rho,T), J(\rho,T)$ can in principle be computed by using formulas (3.18), (4.4), (3.23). Namely, $H(T[\rho])$ is given by formula (3.18) with α replaced by α' computed via (4.4), and

$$H(\rho,T)=\frac{1}{2}\,\mathrm{Sp}\,g\left(\mathrm{abs}(\Delta_{12}^{-1}\alpha_{12}')-\frac{I}{2}\right),$$

where

$$\alpha_{12}'=\begin{bmatrix}\alpha' & K\beta\\ \beta^T K^T & \alpha\end{bmatrix},$$

$$\beta=\Delta\sqrt{-(\Delta^{-1}\alpha)^2-I/4}\,,$$

is computed by inserting (4.3) into

$$\alpha_{12}'-\frac{i}{2}\Delta_{12}'=\mathrm{Tr}\begin{bmatrix}R'\\ R_2\end{bmatrix}\rho[R'^T,R_2^T],$$

where R_2 are the (unchanged) canonical observables of the reference system.

Alternatively, the entropy exchange can be calculated as the output entropy $H(T_E[\rho])$ if an explicit description of T_E is available. We shall demonstrate this method in the example of one-mode channels in Appendix C.

D. Norms of Gaussian transformations

The transposition operation on a bosonic system can be realized as the time-reversal operation, i.e., the operation reversing the signs of all momentum operators, while leaving the position operators unchanged. Obviously, the dual T^* then takes Weyl operators into Weyl operators. So transposition is just like a linear bosonic channel, albeit without the scalar factor $f(z')$ in Eq. (4.2). It is this factor which makes the difference between positivity and complete positivity, and also enters the norm $\|T\|_{\mathrm{cb}}$. In this section we will provide general criteria for deciding complete positivity and computing the norm of general linear bosonic transformations.

These are by definition the operators T acting on Weyl operators according to Eq. (4.1) where $f(z')$ is a scalar factor. We will assume for simplicity (and in view of the applications in the following sections) that the antisymmetric form

$$\Delta''(z_1,z_2)=\Delta'(z_1,z_2)-\Delta(K^Tz_1,K^Tz_2) \tag{4.8}$$

is nondegenerate. This makes the space Z' with the form Δ'' into a phase space in its own right. With the introduction of suitable canonical coordinates it becomes isomorphic to (Z,Δ), so there exists an invertible linear operator $A:Z\rightarrow Z$ such that $\Delta''(z_1,z_2)=\Delta(A^{-1}z_1,A^{-1}z_2)$.

If f is continuous and has sufficient decay properties (which will be satisfied in our applications), there is a unique trace class operator ρ determined by the equation

$$\mathrm{Tr}(\rho V(z))=f(Az). \tag{4.9}$$

Then T is completely positive if and only if ρ is a positive trace class operator. This is a standard result in the theory of quasifree maps on CCR algebras [9]. It is proved by showing that both properties are equivalent to a "twisted positive definiteness condition," namely the positive definiteness of all matrices of the form

$$M_{rs}=f(z_r-z_s)\exp\left(-\frac{i}{2}\Delta'(z_r,z_s)+\frac{i}{2}\Delta(K^Tz_r,K^Tz_s)\right),$$

where $z_1,\ldots,z_n$ are an arbitrary choice of n phase-space points.

If ρ is a nonpositive Hermitian trace class operator, it has a unique decomposition into the positive and negative part: $\rho=\rho_+-\rho_-$ such that $\rho_\pm\geqslant 0$, and $\rho_+\rho_-=0$. Then $|\rho|=\rho_++\rho_-$ and the trace norm is $\|\rho\|_1=\mathrm{Tr}(\rho_+)+\mathrm{Tr}(\rho_-)$. Inserting $\rho_\pm$ into Eq. (4.9) instead of ρ, we get two functions $f_\pm$ on phase space and from Eq. (4.1) two linear bosonic transformations $T_\pm$ with $T=T_+-T_-$. By the criterion just proved, T_+ and T_- are completely positive. Hence

$$\|T\|_{\text{cb}} \leq \|T_+\|_{\text{cb}} + \|T_-\|_{\text{cb}} = \|T_+^*[I]\| + \|T_-^*[I]\| \\ = f_+(0) + f_-(0) = \text{Tr}(\rho_+) + \text{Tr}(\rho_-) = \|\rho\|_1 . \tag{4.10}$$

If the factor f is a Gaussian, i.e.,

$$f(z) = \exp(-\tfrac{1}{2} z^T \beta z) \tag{4.11}$$

for some positive definite matrix β, we can go one step further. In this case we may decompose β into normal modes with respect to Δ'', which decomposes T into a tensor product of one-mode Gaussian transformations T_{ℓ}, for each of which $\|T_{\ell}\|_{\text{cb}}$ may be computed separately by the above method. This amounts to computing the trace norm of the operator ρ_γ given by Eq. (3.15) with arbitrary positive γ. The absolute value of ρ_γ is obtained by taking absolute values of all the eigenvalues, which still makes $\|\rho_\gamma\|_1$ a geometric series,

$$\|\rho_\gamma\|_1 = \frac{1}{\gamma + 1/2} \sum_{n=0}^{\infty} \left| \frac{\gamma - 1/2}{\gamma + 1/2} \right|^n = \max\left\{ 1, \frac{1}{2\gamma} \right\}. \tag{4.12}$$

This is all the information we need for the estimates of quantum capacity in the following section.

V. THE CASE OF ONE MODE

A. Attenuation and amplification channels with classical noise

The channel we consider in this section combines attenuation or amplification [14] with additive classical noise [18]. It can also be described as the most general one-mode gauge invariant channel, or in quantum optics terminology, the most general one-mode channel not involving squeezing. Channels of this type were also used in [28] as the basis for an analysis of the classical limit of quantum mechanics.

Let us consider the CCR with one degree of freedom $a = (1/\sqrt{2\hbar})(q + ip)$, and let a_0 be another mode in the Hilbert space $\mathcal{H}_0 = \mathcal{H}_E$ of an ‘‘environment.’’ Let the environment be initially in the vacuum state, i.e., in the state with the characteristic function (3.14) with $N = 0$. Let ξ be a complex random variable with zero mean and variance N_c describing additive classical noise in the channel. The linear attenuator with coefficient $k < 1$ and the noise N_c is described by the transformation

$$a' = ka + \sqrt{1 - k^2} a_0 + \xi$$

in the Heisenberg picture. Similarly, the linear amplifier with coefficient $k > 1$ is described by the transformation

$$a' = ka + \sqrt{k^2 - 1} a_0^\dagger + \xi.$$

It follows that the corresponding transformations $T[\rho]$ of states in the Schrödinger picture both have the characteristic function

$$\text{Tr}\, T[\rho] V(z) = \text{Tr}\, \rho V(kz) \exp\left[-\frac{\hbar}{2} (|k^2 - 1|/2 + N_c) |z|^2 \right]. \tag{5.1}$$

Let the input state ρ of the system be the elementary Gaussian with the characteristic function (3.14). Then the entropy of ρ is $H(\rho) = g(N)$. From Eq. (5.1) we find that the output state $T[\rho]$ is again elementary Gaussian with N replaced by

$$N' = k^2 N + N_0',$$

where

$$N_0' = \max\{0, (k^2 - 1)\} + N_c$$

is the value of the output mean photon number corresponding to the input vacuum state. Then

$$H(T[\rho]) = g(N'). \tag{5.2}$$

Now we calculate the exchange entropy $H(\rho, T)$. The (pure) input state ρ_{12} of the extended system $\mathcal{H}_1 \otimes \mathcal{H}_2$ is characterized by the 2×2 matrix (3.24). The action of the extended channel $(T \otimes \text{id})$ transforms this matrix into

$$\Delta_{12}^{-1} \tilde{\alpha}_{12} \leftrightarrow \begin{bmatrix} i(N' + \frac{1}{2}) & k\sqrt{N(N+1)} \\ k\sqrt{N(N+1)} & -i(N + \frac{1}{2}) \end{bmatrix}.$$

From formula (3.17) we deduce $H(\rho, T) = g(|\lambda_1| - \frac{1}{2}) + g(|\lambda_2| - \frac{1}{2})$, where λ_1, λ_2 are the eigenvalues of the complex matrix in the right-hand side. Solving the characteristic equation we obtain

$$\lambda_{1,2} = \frac{i}{2} [(N' - N) \pm D], \tag{5.3}$$

where $D = \sqrt{(N + N' + 1)^2 - 4k^2 N(N+1)}$. Hence

$$H(\rho, T) = g\left(\frac{D + N' - N - 1}{2} \right) + g\left(\frac{D - N' + N - 1}{2} \right). \tag{5.4}$$

Now using the theorem of Sec. V, we can calculate the quantity

$$C_e(T) = I(\rho, T) = H(\rho) + H(T[\rho]) - H(\rho, T)$$

as a function of the parameters N, k, N_c, and try to compare it with the one-shot unassisted classical capacity of the channel $C_1(T)$ given by expression (2.4) where the maximum is taken over all probability distributions $\{p_i\}$ and the collections of density operators $\{\rho_i\}$, satisfying the power constraint $\Sigma_i p_i \,\text{Tr}\, \rho_i a^\dagger a \leq N$. It is quite plausible, but not yet proven that this maximum is achieved on coherent states with the Gaussian probability density $p(z) = (\pi N)^{-1} \exp(-|z|^2/N)$, giving the value

$$\underline{C}_1(T) = g(N') - g(N_0').$$

The ratio

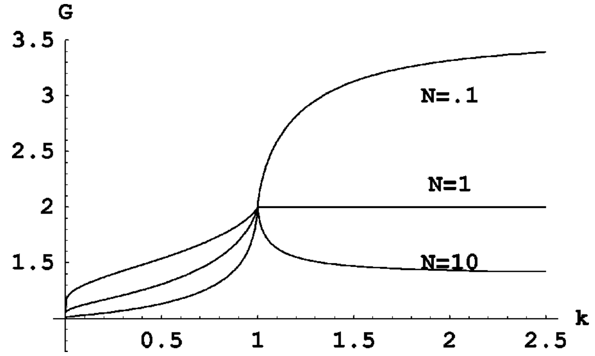

FIG. 1. Gain of entanglement assistance. Gain (5.5) as a function of k with $N_c=0$. Parameter=input noise N.

$$G=\frac{C_e}{\underline{C}_1}\tag{5.5}$$

then gives at least an upper bound for the *gain* of using entanglement-assisted versus unassisted classical capacity. In particular, when the signal mean photon number N tends to zero while $N_0'>0$,

$$\underline{C}_1(T)\sim Nk^2\log\left(\frac{N_0'+1}{N_0'}\right),$$

$$C_e(T)\sim -N\log N/(N_0'+1),$$

and G tends to infinity as $-\log N$.

The plots of G as function of k for $N_c=0$, and as a function of N_c for $k=1$ are given in Figs. 1 and 2, respectively. The behavior of the entropies $H(T[\rho]),H(\rho,T)$ as functions of k for $N_c=0$ is clear from Fig. 3. For all N the coherent information $H(T[\rho])-H(\rho,T)$ turns out to be positive for $k>1/\sqrt{2}$ and negative otherwise. It tends to $-H(\rho)$ for $k\to 0$, is equal to $H(\rho)$ for $k=1$, and quickly tends to zero as $k\to\infty$ (see Fig. 4).

B. Estimating the quantum capacity

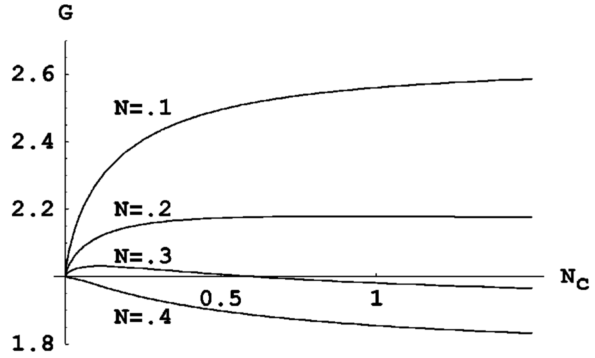

FIG. 2. Gain of entanglement assistance. Gain (5.5) as a function of N_c with $k=1$. Parameter=input noise N.

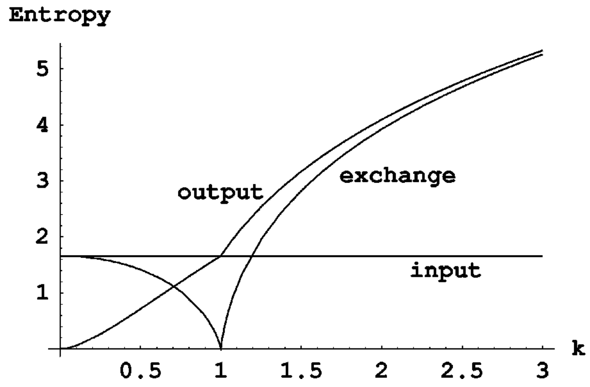

FIG. 3. Entropies. Output entropy from Eq. (5.2); exchange entropy from Eq. (5.4) with $N_c=0$.

Going back to the upper bound for quantum capacity in Sec. IV, we see that T is given by Eq. (4.1) with $Kz=kz$ and

$$f(z)=\exp\left(-\frac{(|k^2-1|/2+N_c)}{2}|z|^2\right).$$

Then $\Delta''=(1-k^2)\Delta$, and the operator A mapping the symplectic form Δ'' to the standard form Δ is multiplication by $\sqrt{|k^2-1|}$, combined for $k>1$ with a mirror reflection to change the sign. This leaves

$$f(Az)=\exp\left(-\frac{(|k^2-1|/2+N_c)}{2|k^2-1|}|z|^2\right),\tag{5.6}$$

i.e., $\rho=\rho_\gamma$ with Eqs. (3.15) and (4.9), where $\gamma=1/2+N_c/|k^2-1|$. This is the verification of the complete positivity of T by the methods of Sec. V A. Of course, this is, strictly speaking, unnecessary, because T was constructed explicitly as a completely positive operator in terms of its dilation in Sec. IV A.

But let us now consider $T\Theta$. It is also a bosonic linear transformation, in which Θ only has the effect of changing the sign of the symplectic form, without changing f. Thus $\Delta''=(1+k^2)\Delta$, and

$$f(Az)=\exp\left(-\frac{(|k^2-1|/2+N_c)}{2|k^2+1|}|z|^2\right).$$

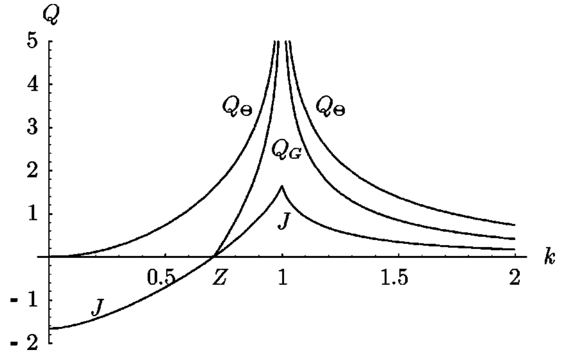

FIG. 4. Bounds for quantum capacity, $N_c=0$. J is the coherent information (5.8) with $N=0.7$; Q_G is the bound maximized over Gaussians (5.9); Q_Θ is the bound on this quantity from transposition (5.7); Z is the zero at $k=1/\sqrt{2}$, common to all curves of type J.

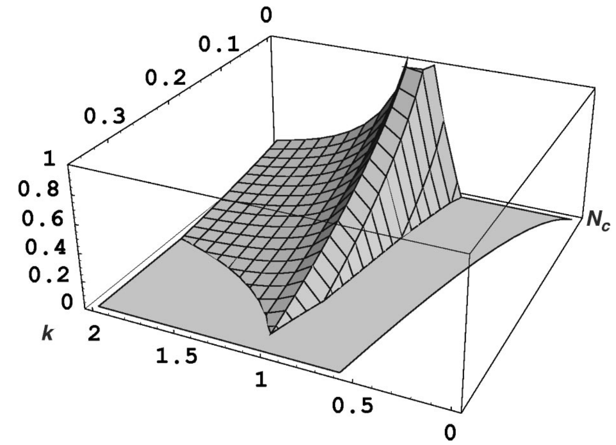

FIG. 5. Gaussian maximized coherent information $Q_G(T)$ as a function of k and N_c. The shaded area is the area, where $Q_\Theta \geq 0$.

which seems like a rather minor change over Eq. (5.6). However, we now get $\rho=\rho_\gamma$ with $\gamma=(|k^2-1|/2+N_c)/(k^2+1)$ which is not necessarily greater than or equal to 1/2, so $T\Theta$ is not necessarily completely positive. Taking the logarithm of Eq. (4.12) we get

$$Q_\Theta(T)\leq\max\{0,\log(k^2+1)-\log(|k^2-1|+2N_c)\}. \tag{5.7}$$

In particular, for $\gamma\geq 1/2$, i.e., for $N_c\geq(|k^2+1|-|k^2-1|)/2 =\max\{1,k^2\}$, the capacities $Q_\Theta(T)$, and hence $Q_\varepsilon(T)$ and $Q(T)$, all vanish.

This upper bound on the quantum capacity is interesting to compare with the quantity $Q_G(T)=\sup J(\rho,T)$, where $J(\rho,T)=H(T[\rho])-H(\rho,T)$, and the supremum is taken over all *Gaussian* input states. Since the coherent information

$$J(\rho,T)=g(N')-g\left(\frac{D+N'-N-1}{2}\right)-g\left(\frac{D-N'+N-1}{2}\right) \tag{5.8}$$

increases with the input power N, we obtain

$$\begin{aligned}Q_G(T)&=\lim_{N\to\infty}J(\rho,T)\\&=\log k^2-\log|k^2-1|-g(N_c/|k^2-1|),\end{aligned} \tag{5.9}$$

which is in a good agreement with the upper bound (5.7) (see Figs. 4 and 5).

ACKNOWLEDGMENTS

A.H. appreciated the illuminating discussion of fragments of the unpublished paper [18] with C. H. Bennett and P. W. Shor. He acknowledges the hospitality of R.F. Werner at the Institute for Mathematical Physics, Technical University of Braunschweig, and he acknowledges the A. von Humboldt Foundation for financial support.

APPENDIX A: PROOF OF PROPERTIES OF Q_Θ

1. Exact additivity

We use the facts that the transpose on a tensor product is the tensor product of the transposes ("$\Theta=\Theta\otimes\Theta$," see above), and the fact that the cb-norm is multiplicative, i.e., $\|T\otimes S\|_{\rm cb}=\|T\|_{\rm cb}\|S\|_{\rm cb}$, for arbitrary linear operators T,S. Hence $\|(T\otimes S)\Theta\|_{\rm cb}=\|(T\Theta)\otimes(S\Theta)\|_{\rm cb}=\|T\Theta\|_{\rm cb}\|S\Theta\|_{\rm cb}$.

2. If T is a channel, so is $\Theta T\Theta$

Indeed, for any n, $(\Theta T\Theta)\otimes{\rm id}_n=(\Theta T\Theta)\otimes(\Theta_n\Theta_n) =(\Theta\otimes\Theta_n)(T\otimes{\rm id}_n)(\Theta\otimes\Theta_n)$ is the product of three positive (although not completely positive) maps. Hence $\Theta T\Theta$ is completely positive. It is also normalized as a channel, because Θ leaves the unit operator and the trace invariant.

3. The bottleneck inequality

We use the inequality $\|TS\|_{\rm cb}\leq\|T\|_{\rm cb}\|S\|_{\rm cb}$, which follows because the cb-norm is defined in terms of an operator norm. Thus $\|TS\Theta\|_{\rm cb}\leq\|T\|_{\rm cb}\|S\Theta\|_{\rm cb}\leq\|S\Theta\|_{\rm cb}$, because $\|T\|_{\rm cb}=1$ for any channel. On the other hand, $\|TS\Theta\|_{\rm cb} =\|(T\Theta)(\Theta S\Theta\|_{\rm cb}\leq\|T\Theta\|_{\rm cb}$, because $\Theta S\Theta$ is also a channel by Appendix A 2. Taking the logarithm of these upper bounds on $\|TS\Theta\|_{\rm cb}$, we find the desired inequality.

4. $Q_\Theta(T)=0$ for separable channels

For a commutative algebra $\mathcal{C}$, Θ is the identity. Therefore, if ${\rm id}_{\mathcal{C}}$ denotes the identity (=ideal channel) on a classical system, we get $\|{\rm id}_{\mathcal{C}}\Theta\|_{\rm cb}=\|{\rm id}_{\mathcal{C}}\|_{\rm cb}=1$, hence $Q_\Theta({\rm id}_{\mathcal{C}})=0$. By the bottleneck inequality a factorization $T =PM$ into preparation and measurement implies $Q_\Theta(T) =Q_\Theta(P\,{\rm id}_{\mathcal{C}}M)\leq\min\{Q_\Theta(P),Q_\Theta({\rm id}_{\mathcal{C}}),Q_\Theta(M)\}=0$. More generally, we find $Q_\Theta(T)=0$ for any "entanglement binding channel," in the sense of [30] which are precisely those channels, for which $T\Theta$ is completely positive.

5. Connection with an entanglement measure

Let

$$E_\Theta(\rho)=\log\|({\rm id}\otimes\Theta)[\rho]\|_1 \tag{A1}$$

denote the entanglement measure mentioned in the text. Using exactly the same techniques as above, one shows that this is a strictly additive upper bound on the distillation rates of pure singlets from ρ. The connection with Q_Θ arises from the problem of estimating the entanglement of a state after one of the subsystems has been sent through a noisy channel T, i.e., the entanglement of a state of the form $(T\otimes{\rm id})[\rho]$. We get

$$\begin{aligned}\|({\rm id}\otimes\Theta)(T\otimes{\rm id})[\rho]\|_1&=\|(T\otimes\Theta)[\rho]\|_1\\&\leq\|T\otimes\Theta\|\,\|\rho\|_1\\&=\|[(T\Theta)\otimes{\rm id}](\Theta\otimes\Theta)\|\\&=\|(T\Theta)\otimes{\rm id}\|\leq\|T\Theta\|_{\rm cb},\end{aligned}$$

and hence

$$E_\Theta((T\otimes\mathrm{id})[\rho])\leqslant Q_\Theta(T). \tag{A2}$$

Moreover, since the operator norm of any Hermiticity preserving operator T can be written as $\|T\|=\sup_\sigma\|T[\sigma]\|_1$, a supremum over pure states, we find that the supremum of the left-hand side in Eq. (A2) over all pure states σ equals the right-hand side. In other words, Q_Θ can be *defined* as the maximal entanglement (as measured by E_Θ) of states transmitted through T.

APPENDIX B: MINIMIZING CONVEX FUNCTIONS OF A DENSITY OPERATOR

There is a useful lemma in classical information theory which gives necessary and sufficient conditions for the global minimum of a convex function of probability distributions in terms of the first partial derivatives [29]. This can be generalized to functions depending on density operators rather than probability distributions.

Let F be a convex function on the set of density operators Σ, and ρ_0 a density operator. In order F to achieve minimum on ρ_0, it is necessary and sufficient that for arbitrary density operator σ the convex function $F((1-t)\rho_0+t\sigma)$ of the real variable t achieves minimum at $t=0$. For this, it is necessary and sufficient that

$$\nabla_X F(\rho_0)\equiv\frac{d}{dt}\Big|_{t=0}F((1-t)\rho_0+t\sigma)\geqslant 0, \tag{B1}$$

where $X=\sigma-\rho_0$, and $\nabla_X F(\rho_0)$ is the directional derivative of F in the direction X, assuming that the derivatives exist. If $\sigma=\Sigma_i p_i\ \sigma_i$, then $\nabla_X F(\rho_0)=\Sigma_i p_i\ \nabla_{X_i}F(\rho_0)$, where $X_i=\sigma_i-\rho_0$. Therefore it is necessary and sufficient that Eq. (B1) holds for pure σ.

If $(1-t)\rho_0+t\sigma\geqslant 0$ for small negative t, then we say that the direction $\overrightarrow{\sigma\rho_0}$ is *inner*. In that case Eq. (B1) takes the form

$$\nabla_X F(\rho_0)=0. \tag{B2}$$

If ρ_0 is nondegenerate, then the direction $\overrightarrow{\sigma\rho_0}$ is inner for arbitrary pure σ in the range of $\sqrt{\rho_0}$, and the necessary and sufficient condition for the minimum is that Eq. (B2) holds for arbitrary such σ.

Let $A_i, i=1,\dots,r$ be a collection of self-adjoint *constraint operators*. Assume that for some real constants λ_i

$$\nabla_X F(\rho_0)=\mathrm{Tr}\,X\sum_i\ \lambda_i A_i\,. \tag{B3}$$

It follows that the convex function $F(\rho)-\mathrm{Tr}\,\rho\Sigma_i\lambda_iA_i$ achieves minimum at the point ρ_0, hence the function $F(\rho)$ achieves minimum at the point ρ_0 under the constraints $\mathrm{Tr}\,\rho A_i=\mathrm{Tr}\,\rho_0A_i,\ \ i=1,\dots,r$.

APPENDIX C: QUANTUM SIGNAL PLUS CLASSICAL NOISE

Let us consider CCR with one degree of freedom described by one mode annihilation operator $a=(1/\sqrt{2\hbar})(q+ip)$, and consider the transformation

$$a'=a+\xi,$$

where ξ is a complex random variable with zero mean and variance N_c. This is a transformation of the type (4.3) with $\Delta_E=0$, which describes the quantum mode in a classical Gaussian environment. The action of the dual channel is

$$T^*[f(a,a^\dagger)]=\int f(a+z,(a+z)^\dagger)\mu_{N_c}(d^2z),$$

where $z=(1/\sqrt{2\hbar})(x+iy)$ is now a complex variable, and $\mu_{N_c}(d^2z)$ is a complex Gaussian probability measure with zero mean and variance N_c, while the channel itself can be described by the formula

$$T[\rho]=\int D(z)\rho D(z)^*\mu_{N_c}(d^2z), \tag{C1}$$

where $D(z)=\exp[i(za^\dagger-\overline{za})]$ is the displacement operator.

The entanglement-assisted classical capacity of the channel (C1) was first studied in [18] by using a rather special way of purification and the computation of the entropy exchange. A general approach following the method of [14] was described in Secs. IV and V; here we give an alternative solution based on the computation of the environment entropy.

For this we need to extend the environment to a quantum system in a pure state. Consider the environment Hilbert space $\mathcal{H}_E=L^2(\mu_{N_c})$ with the vector $|\Psi_0\rangle$ given by the function identically equal to 1. The tensor product $\mathcal{H}\otimes\mathcal{H}_E$ can be realized as the space $L^2_{\mathcal{H}}(\mu_{N_c})$ of μ_{N_c}-square integrable functions $\psi(z)$ with values in $\mathcal{H}$. Define the unitary operator U in $\mathcal{H}\otimes\mathcal{H}_E$ by

$$(U\psi)(z)=D(z)\psi(z).$$

Then

$$T[\rho]=\mathrm{Tr}_{\mathcal{H}_E}U(\rho\otimes|\Psi_0\rangle\langle\Psi_0|)U^*,$$

while

$$T_E[\rho]=\mathrm{Tr}_{\mathcal{H}}\,U(\rho\otimes|\Psi_0\rangle\langle\Psi_0|)U^*.$$

This means that $T_E[\rho]$ is an integral operator in $L^2(\mu_{N_c})$ with the kernel

$$K(z,z')=\mathrm{Tr}\,D(z)\rho_0D(z')^*$$
$$=\exp[i\,\mathrm{Im}\,\overline{z'}z-(E+1/2)|z-z'|^2].$$

Let us define the unitary operators $V(z_1,z_2)$ in $L^2(\mu_{N_c})$ by

$$V(z_1,z_2)\psi(z)=\psi(z+z_2)\exp\left[i\,\mathrm{Re}\,\overline{z_1}\left(z+\frac{z_2}{2}\right)-\frac{1}{N_c}\,\mathrm{Re}\,\overline{z_2}\left(z+\frac{z_2}{2}\right)\right].$$

The operators $V(z_1,z_2)$ satisfy the Weyl-Segal CCR with two degrees of freedom with respect to the symplectic form

$$\Delta((z_1,z_2),(z_1',z_2'))=\mathrm{Re}(\overline{z_1'}z_2-\overline{z_1}z_2').$$

Passing over to the real variables x,y one finds the corresponding commutation matrix

$$\Delta_E=\hbar\begin{bmatrix}0&0&-1&0\\0&0&0&-1\\1&0&0&0\\0&1&0&0\end{bmatrix}.$$

The characteristic function of the operator $T_E[\rho_0]$ is

$$\mathrm{Tr}\,T_E[\rho_0]V(z_1,z_2)=\int V(z_1,z_2)K(\check{z},z)|_{\check{z}=z}\mu_{N_c}(d^2z),$$

where $V(z_1,z_2)$ acts on K as a function of the argument $\check{z}$. Evaluating the Gaussian integral, we obtain that it is equal to

$$\exp\left[-\frac{1}{4}\left(N_c|z_1|^2+2N_c\mathrm{Im}\overline{z_1}z_2+\frac{D^2}{N_c}|z_2|^2\right)\right],$$

(where now $D=\sqrt{(N_c+1)^2+4N_cN}$), which is a Gaussian characteristic function with the correlation matrix

$$\alpha_E'=\frac{\hbar}{2}\begin{bmatrix}N_c&0&0&N_c\\0&N_c&-N_c&0\\0&-N_c&\dfrac{D^2}{N_c}&0\\N_c&0&0&\dfrac{D^2}{N_c}\end{bmatrix}.$$

Thus

$$\Delta_E^{-1}\alpha_E'=\frac{1}{2}\begin{bmatrix}0&-N_c&\dfrac{D^2}{N_c}&0\\N_c&0&0&\dfrac{D^2}{N_c}\\-N_c&0&0&-N_c\\0&-N_c&N_c&0\end{bmatrix}.$$

By using the Pauli matrix σ_y, we can write it as

$$\frac{1}{2}\begin{bmatrix}-i\sigma_yN_c&\dfrac{D^2}{N_c}\\-N_c&-i\sigma_yN_c\end{bmatrix}=\frac{1}{2}\begin{bmatrix}I&0\\0&\sigma_y\end{bmatrix}\begin{bmatrix}-i\sigma_yN_c&\sigma_y\dfrac{D^2}{N_c}\\-\sigma_yN_c&-i\sigma_yN_c\end{bmatrix}\begin{bmatrix}I&0\\0&\sigma_y\end{bmatrix},$$

hence the absolute values of the eigenvalues of $\Delta_E^{-1}\alpha_E'$ are the same as that of the matrix

$$\begin{bmatrix}iN_c&-\dfrac{D^2}{N_c}\\N_c&iN_c\end{bmatrix},$$

which coincide with Eq. (5.3) in the case $k=1$.

[1] C. H. Bennett and P. W. Shor, IEEE Trans. Inf. Theory **IT-44**, 2724 (1998).

[2] A. S. Holevo, Russian Math. Surveys **53**, 1295 (1998); e-print quant-ph/9809023 (unpublished).

[3] H. Barnum, M. A. Nielsen, and B. Schumacher, Phys. Rev. A **57**, 4153 (1998); e-print quant-ph/9702049 (unpublished).

[4] C. Adami and N. J. Cerf, Phys. Rev. A **56**, 3470 (1997); e-print quant-ph/9609024 (unpublished).

[5] H. Barnum, E. Knill, and M. A. Nielsen, IEEE Trans. Inform. Theory **46**, 1317 (2000); e-print quant-ph/9809 (unpublished).

[6] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).

[7] S. L. Braunstein, e-print quant-ph/9904002 (unpublished).

[8] A. Furusawa, J. Sϕrensen, S. L. Braunstein, C. Fuchs, H. J. Kimble, and E. S. Polzik, Science **282**, 706 (1998).

[9] B. Demoen, P. Vanheuverzwijn, and A. Verbeure, Rep. Math. Phys. **15**, 27 (1979).

[10] A. S. Holevo, IEEE Trans. Inf. Theory **IT-21**, 533 (1975).

[11] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982), Chap. 5.

[12] R. F. Werner, J. Math. Phys. **25**, 1404 (1984).

[13] A. S. Holevo, M. Sohma, and O. Hirota, Phys. Rev. A **59**, 1820 (1998).

[14] A. S. Holevo, in *Quantum Communication, Computation, and Measurement,* edited by P. Kumar, M. D'Ariano, and O. Hirota (Plenum, New York, 1999), pp. 75–82; e-print quant-ph/9809022 (unpublished).

[15] A. S. Holevo, Probl. Inf. Transm. **8**, 63 (1972).

[16] K. Kraus, *States, Effects and Operations*, Lecture Notes in Physics Vol. 190 (Springer, Berlin, 1983).

[17] G. Lindblad, *Quantum Aspects of Optical Communication*, edited by C. Benjaballah, O. Hirota, and S. Reynaud, Lecture Notes in Physics Vol. 378 (Springer, Berlin, 1991), pp. 71–80.

[18] P. W. Shor *et al.* (unpublished).

[19] S. Lloyd, Phys. Rev. A **56**, 1613 (1997).

[20] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Phys. Rev. A **57**, 830 (1998).

[21] V. I. Paulsen, *Completely Bounded Maps and Dilations* (Longman Scientific and Technical, New York, 1986).

[22] D. Aharonov, A. Kitaev, and N. Nisan, e-print quant-ph/9806029 (unpublished).

[23] T. Ogawa and H. Nagaoka, IEEE Trans. Inf. Theory **45**, 2486 (1999); e-print quant-ph/9808063 (unpublished).

[24] A. Winter, IEEE Trans. Inf. Theory **45**, 2481 (1999).

[25] R. Simon, M. Selvadoray, and G. S. Agarwal, Phys. Rev. (to be published).

[26] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976), Chap. 5.

[27] A. S. Holevo, Theor. Math. Phys. **6**, 3–20 (1971).

[28] R. F. Werner, e-print quant-ph/9504016 (unpublished).

[29] R. Gallager, *Information Theory and Reliable Communications* (Wiley, New York, 1968), Chap. 4.

[30] P. Horodecki, M. Horodecki, and R. Horodecki, Phys. Rev. A **61**, 052309 (2000); e-print quant-ph/9905058.