# Capacity of quantum Gaussian channels

A. S. Holevo
*Steklov Mathematical Institute, Moscow, Russia*

M. Sohma
*Matsushita Research Institute, Incorporated, Tokyo, Japan*

O. Hirota
*Research Center for Quantum Communications, Tamagawa University, Tokyo, Japan*
(Received 11 August 1998)

The aim of this paper is to give explicit calculation of the classical capacity of quantum Gaussian channels, in particular, involving squeezed states. The calculation is based on a general formula for the entropy of a quantum Gaussian state, which is of independent interest, and on the recently proved coding theorem for quantum communication channels. [S1050-2947(99)00802-1]

PACS number(s): 03.67.−a, 03.65.Bz, 42.50.Dv, 89.70.+c

## I. INTRODUCTION

The question of fundamental physical limitations on the quality and rate of information transmission is at the core of quantum theory of communication, which has led to profound and exciting new insights into both physics and information transmission science. This theory provides a general framework for study of communication processes, in which classical information is conveyed by quantum states. One of the recent achievements of the quantum information theory is the direct coding theorem for transmission of classical information through quantum communication channels [1–3], which provides an explicit formula for the capacity of the channel as supremum of the quantum entropy bound with respect to input probability distributions. This result was recently extended to channels with constrained inputs [4] among which the channels with additive quantum Gaussian noise and the constrained power of the signal are the most important. The aim of the present paper is a further study of this case, in particular, explicit calculation of the classical capacity of the memoryless squeezed state channel (Sec. IV B). This allows us to quantitatively evaluate the properties of squeezed states from an information-theoretic point of view.

The core of our calculations, apart from the above mentioned coding theorem, is the formula for the von Neumann entropy of the general quantum Gaussian state (Sec. IV A), which may have independent interest. The natural class of quantum Gaussian states which includes, in particular, coherent and squeezed states, as well as their thermal mixtures is discussed in Sec. III B. This is preceded by a description of the more familiar gauge-invariant Gaussian states (those having complex Gaussian $P$ representation in terms of coherent states), which we include in Sec. III A for completeness.

## II. CLASSICAL SIGNAL PLUS QUANTUM NOISE

Let us describe the process of classical information transmission through a quantum memoryless channel with a continuous signal parameter in the model of optical communications. Consider a quantum system, such as a cavity field with finite numbers of modes [5], described by annihilation operators $a_1, \ldots, a_s$ satisfying canonical commutation relation (CCR)

$$[a_j, a_k^\dagger] = \delta_{j,k} I, \quad [a_j, a_k] = 0. \tag{1}$$

Let $\mathcal{H}$ be the Hilbert space of irreducible representation of CCR (1), and let $\rho(0)$ be a density operator in $\mathcal{H}$ describing the state of the cavity field. Consider the family of density operators

$$\rho(\mu) = \mathcal{D}(\mu)\rho(0)\mathcal{D}(\mu)^\dagger, \quad \mu \in \mathbf{C}^s, \tag{2}$$

where $\mathbf{C}^s$ is an $s$-dimensional complex vector space, and

$$\mathcal{D}(\mu) = \exp\sum_{j=1}^{s} (\mu_j a_j^\dagger - \overline{\mu_j} a_j) \quad \text{for } \mu = (\mu_j) \tag{3}$$

is the unitary *displacement operator* in $\mathcal{H}$.

In an optical communication system $\rho(0)$ describes background noise, comprising quantum noise, and $\mu$ is the classical signal. Thus the mapping $\mu \to \rho(\mu)$ is a classical-quantum (CQ) channel in the sense of [4]. We treat the product memoryless channel in the Hilbert space $\mathcal{H}^{\otimes n} = \mathcal{H} \otimes \cdots \otimes \mathcal{H}$ ($n$ copies), then the signal will be represented by the sequence of vectors $\mu(t)$; $t = 1, \ldots, n$, and the channel mapping is $\mu(\cdot) \to \rho(\mu(1)) \otimes \cdots \otimes \rho(\mu(n))$. We impose the power constraint on the signal:

$$\sum_{t=1}^{n} \left( \sum_{j=1}^{s} \hbar\omega_j |\mu_j(t)|^2 \right) \leq nE, \tag{4}$$

where $\omega_j$ are the frequencies of the modes.

Of course, the memoryless channel is an idealization and should be considered as the first necessary step towards a more realistic model of the ''waveform'' channel, which explicitly takes into account internal dynamics of the field (cf. [4]). According to Theorem 3 of [4], the capacity of such a channel is equal to

 

$$C=\sup_{P\in\mathcal{P}_1}\left(H(\rho_P)-\int H(\rho(\mu))P(d\mu)\right), \tag{5}$$

where $H=-\mathrm{Tr}\rho\ln\rho$ is the von Neumann entropy, $\rho_P=\int\rho(\mu)P(d\mu)$, and $\mathcal{P}_1$ is a convex subset of probability distributions $P(d\mu)$ on $\mathbf{C}^s$, satisfying the inequality

$$\int\sum_{j=1}^{s}\hbar\omega_j|\mu_j|^2P(d\mu)\le E. \tag{6}$$

In the case (2) the operators $\rho(\mu)$ and $\rho(0)$ are unitarily equivalent and we have important simplification $H(\rho(\mu))\equiv H(\rho(0))$, resulting in

$$C=\sup_{P\in\mathcal{P}_1}H(\rho_P)-H(\rho(0)). \tag{7}$$

Of special interest is the case of Gaussian noise. The definition and properties of the general quantum Gaussian density operator were given in [6,7] in terms of symplectic spaces, and will be repeated in the next section in somewhat different vector notations. Let us show here that in the case of Gaussian $\rho(0)$, we can restrict optimization in Eq. (7) to Gaussian probability distribution $P$. For this we need only the two following properties (see the Appendix).

(i) Among all density operators $\rho$ in $\mathcal{H}$ with fixed first and second moments

$$\mathrm{Tr}\,\rho a_j,\quad \mathrm{Tr}\,\rho a_j^\dagger a_k,\quad \mathrm{Tr}\,\rho a_j a_k \tag{8}$$

the Gaussian density operator is one which has maximal entropy. This fundamental property explains, in particular, the importance of the Gaussian states from the viewpoint of the Jaynes ''maximum ignorance'' principle.

(ii) The mixture $\rho_P$ of the Gaussian density operators $\rho(\mu)$ with Gaussian probability distribution $P(d\mu)$ is a Gaussian density operator.

Now for any $P\in\mathcal{P}_1$ let $\tilde{P}$ be the Gaussian probability distribution with the same first and second moments

$$\int\mu_jP(d\mu),\quad \int\overline{\mu_j}\mu_kP(d\mu),\quad \int\mu_j\mu_kP(d\mu), \tag{9}$$

then (1) $\tilde{P}\in\mathcal{P}_1$ because Eq. (6) involves only second moments; (2) by (ii) $\rho_{\tilde{P}}$ is Gaussian; and (3) $\rho_P$ and $\rho_{\tilde{P}}$ have the same second moments. Indeed, for any polynomial $F(\mu,\overline{\mu})$ of second order

$$\begin{aligned}\mathrm{Tr}\,\rho_PF(a,a^\dagger)&=\int\mathrm{Tr}\,\rho(\mu)F(a,a^\dagger)P(d\mu)\\&=\int\mathrm{Tr}\,\rho(0)\mathcal{D}(\mu)^\dagger F(a,a^\dagger)\mathcal{D}(\mu)P(d\mu)\\&=\mathrm{Tr}\,\rho(0)\int F(a-\mu,(a-\mu)^\dagger)P(d\mu)\\&=\mathrm{Tr}\,\rho(0)\int F(a-\mu,(a-\mu)^\dagger)\tilde{P}(d\mu)\\&=\mathrm{Tr}\,\rho_{\tilde{P}}F(a,a^\dagger).\end{aligned} \tag{10}$$

Thus if $\rho(0)$ is Gaussian, then in Eq. (7) we can restrict ourselves to Gaussian $P$, and by (ii) $\rho_P$ will be a Gaussian density operator. Therefore to calculate Eq. (7) we need to know the entropy of a general Gaussian density operator.

## III. GAUSSIAN DENSITY OPERATORS

As mentioned in the preceding section, the case of Gaussian quantum noise is most important. In this section we recall the general definition and properties of quantum Gaussian density operators to be applied later in the calculation of the channel capacity.

### A. Gauge-invariant case

We first consider the density operator which has Glauber's $P$ representation

$$\rho(0)=\pi^{-s}|\det N|^{-1}\int\exp(-\zeta^\dagger N^{-1}\zeta)|\zeta\rangle\langle\zeta|d^{2s}\zeta \tag{11}$$

(see [5], Chap. V, Sec. 5. II). Here $\zeta\in\mathbf{C}^s$, $|\zeta\rangle$ are the coherent vectors in $\mathcal{H}$, $a|\zeta\rangle=\zeta|\zeta\rangle$, and $N$ is a positive Hermitian matrix such that

$$N=\mathrm{Tr}\,a\rho a^\dagger \tag{12}$$

(we use here vector notations, where $a=[a_1,\dots,a_s]^T$ is a column vector and $a^\dagger=[a_1^\dagger,\dots,a_s^\dagger]$ is a row vector). Such density operators, respecting the complex structure associated with the coherent states, will be called *gauge-invariant*, because they are invariant with respect to the gauge transformation of the first kind: $a\to e^{i\varphi}a$. There exists a unitary operator $U$ in $\mathbf{C}^s$ such that

$$UNU^*=\begin{bmatrix}N_1 & & 0\\ & \ddots & \\ 0 & & N_s\end{bmatrix}\equiv\mathrm{diag}(N_j). \tag{13}$$

By considering the canonical transformation $a\to\tilde{a}=Ua$, $\tilde{\zeta}=U\zeta$, $|\tilde{\zeta}\rangle^\sim=|\zeta\rangle$, we see that the density operator (11) is decomposed into the tensor product

$$\rho(0)=\bigotimes_{j=1}^{s}\rho_j(0), \tag{14}$$

where

$$\rho_j(0)=\pi^{-1}N_j^{-1}\int\exp\left(-\frac{|\tilde{\zeta}_j|^2}{N_j}\right)|\tilde{\zeta}_j\rangle^{\sim\sim}\langle\tilde{\zeta}_j|d^2\tilde{\zeta}_j, \tag{15}$$

and $H(\rho(0))=\Sigma_{j=1}^sH(\rho_j(0))$. The spectral decomposition of $\rho_j(0)$ is well known (see, e.g., [7])

$$\rho_j(0)=\frac{1}{N_j+1}\sum_{m=0}^{\infty}\left(\frac{N_j}{N_j+1}\right)^m|m\rangle^{\sim\sim}\langle m|, \tag{16}$$

where $|m\rangle^\sim$; $m=0,1,\dots$ are the orthonormal vectors, such that $\tilde{a}_j^\dagger\tilde{a}_j|m\rangle^\sim=m|m\rangle^\sim$. From Eq. (16) $H(\rho_j(0))=g(N_j)$, where

$$g(x)=(x+1)\ln(x+1)-x\ln x,\quad x>0 \tag{17}$$

is a monotonously increasing concave function. Therefore we obtain the well-known formula

$$H(\rho(0))=\sum_{j=1}^{s} g(N_j)=\mathrm{Sp}\, g(\mathrm{diag}(N_j))=\mathrm{Sp}\, g(N), \tag{18}$$

where Sp denotes trace of matrices as distinct from trace of operators Tr.

### B. The general case

The density operator (11) is not the most general which naturally can be called quantum Gaussian. A practical example, not covered by Eq. (11), is the squeezed state, which has no $P$ representation with positive probability density. To explain the general definition, let us change from complex to real setting by introducing canonical pairs

$$q_j=\sqrt{\frac{\hbar}{2\omega_j}}(a_j+a_j^\dagger),\quad p_j=i\sqrt{\frac{\hbar\omega_j}{2}}(a_j^\dagger-a_j), \tag{19}$$

such that

$$a_j=\frac{1}{\sqrt{2\hbar\omega}}(\omega_j q_j+ip_j), \tag{20}$$

satisfying the Heisenberg CCR

$$[q_j,p_k]=i\delta_{jk}\hbar I,\quad [q_j,q_k]=0,\quad [p_j,p_k]=0. \tag{21}$$

Let us introduce the column vector

$$R=[q_1,\dots,q_s;p_1,\dots,p_s]^T. \tag{22}$$

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

$$V(z)=\exp i\sum_{j=1}^{s}(x_jq_j+y_jp_j)=\exp(iR^Tz). \tag{23}$$

The operators $V(z)$ satisfy the Weyl-Segal CCR

$$V(z)V(z')=\exp\left(\frac{i}{2}\Delta(z,z')\right)V(z+z'), \tag{24}$$

where

$$\Delta(z,z')=\hbar\sum_{j=1}^{s}(x_j'y_j-x_jy_j') \tag{25}$$

is the canonical symplectic form. The Weyl-Segal CCR is the rigorous counterpart of the Heisenberg CCR, involving only bounded operators. Let us mention that if $\zeta_k=(1/\sqrt{2\hbar\omega_k})(\omega_kx_k+iy_k)$, then

$$\mathcal{D}(\zeta)=\exp\frac{i}{\hbar}\sum_{j=1}^{s}(y_jq_j-x_jp_j)=V(-\Delta^{-1}z), \tag{26}$$

where

$$\Delta=\left[\begin{array}{cccc|cccc} 0 & & & & \hbar & & & 0 \\ & 0 & & & & \hbar & & \\ & & \ddots & & & & \ddots & \\ & & & 0 & 0 & & & \hbar \\ \hline -\hbar & & & 0 & 0 & & & \\ & -\hbar & & & & 0 & & \\ & & \ddots & & & & \ddots & \\ 0 & & & -\hbar & & & & 0 \end{array}\right] \tag{27}$$

is the $(2s)\times(2s)$-skew-symmetric *commutation matrix* of components of the vector $R$. Most of the results below are valid for the case where the commutation matrix is an arbitrary nondegenerated skew-symmetric matrix, not necessarily of the canonical form (27).

*Definition. The density operator $\rho$ is called Gaussian, if its quantum characteristic function has the form*

$$\mathrm{Tr}\,\rho V(z)=\exp(im^Tz-\tfrac{1}{2}z^T\alpha z), \tag{28}$$

*where m is a column (2s) vector and $\alpha$ is a real symmetric $(2s)\times(2s)$ matrix.*

One can show that

$$m=\mathrm{Tr}\,\rho R,\quad \alpha-\frac{i}{2}\Delta=\mathrm{Tr}\,R\rho R^T \tag{29}$$

(cf. [6,7]). The *mean m* can be an arbitrary vector; the necessary and sufficient condition on the *correlation matrix $\alpha$* is the generalized Robertson uncertainty relation

$$\alpha-\frac{i}{2}\Delta\geqslant 0. \tag{30}$$

This condition is equivalent to its transpose $\alpha+(i/2)\Delta\geqslant 0$, and to the following matrix generalization of the Heisenberg uncertainty relation:

$$\Delta^{-1}\alpha\Delta^{-1}+\frac{1}{4}\alpha^{-1}\geqslant 0, \tag{31}$$

which is obtained by combining Eq. (30) and its transpose. The state $\rho$ is pure if and only if the equality holds in this equation, or

$$(\Delta^{-1}\alpha)^2=-\frac{1}{4}I. \tag{32}$$

This is equivalent to $\det(2\Delta^{-1}\alpha)=1$, and, for the canonical form (27) of the commutation matrix, to the condition

$$\det(2\alpha)=\hbar^{2s} \tag{33}$$

(see the Appendix).

In addition to the two properties listed in Sec. II, we note the following two properties.

(iii) The Gaussian density operator with the correlation matrix $\alpha$ is pure if and only if Eq. (33) holds.

Since Eq. (30) is also a necessary condition for the correlation matrix of the arbitrary density operator, we have the following.

(iv) For any density operator $\rho$ there is Gaussian density operator $\tilde{\rho}$ with the same mean $m$ and correlation matrix $\alpha$.

It is convenient to write $m$ and $\alpha$ in the block form

$$m=[m_1^q,\ldots,m_s^q;m_1^p,\ldots,m_s^q]^T$$
$$\alpha=\left[\begin{array}{c|c}\alpha^{qq} & \alpha^{qp}\\ \hline \alpha^{pq} & \alpha^{pp}\end{array}\right].$$

By using Eq. (26), it can be shown ([7], Sec. V. 5) that

$$\rho(m)=\mathcal{D}(\mu)\rho(0)\mathcal{D}(\mu)^\dagger,\quad \mu_j=\frac{1}{\sqrt{2\hbar\omega_j}}[\omega_j m_j^q+m_j^p],\quad (34)$$

where $\rho(m)$ is the Gaussian density operator with mean $m$ and fixed correlation matrix $\alpha$. This allows us to restrict to the case $m=0$ in calculation of the entropy.

Let us show that the gauge-invariant state (11) is Gaussian in the sense of our definition. The characteristic function is

$$\operatorname{Tr}\rho(0)V(z)=\pi^{-s}|\det N|^{-1}\times\int\exp(-\zeta^\dagger N^{-1}\zeta)\langle\zeta|V(z)|\zeta\rangle d^{2s}\zeta.\quad (35)$$

By using Eq. (26) and matrix elements of the displacement operator [[5], Eq. (3.22) p. 131], we calculate the Gaussian integral in the right-hand side as $\exp(-\frac{1}{2}z^T\alpha z)$, with

$$\alpha=\hbar\Omega\left[\begin{array}{c|c}\mathrm{Re}N+I/2 & -\mathrm{Im}N\\ \hline \mathrm{Im}N & \mathrm{Re}N+I/2\end{array}\right]\Omega,\quad (36)$$

where

$$\Omega=\left[\begin{array}{ccc|ccc}\sqrt{\omega_1}^{-1} & & 0 & 0 & & \\ & \ddots & & & \ddots & \\ 0 & & \sqrt{\omega_s}^{-1} & & & 0\\ \hline 0 & & & \sqrt{\omega_1} & & 0\\ & \ddots & & & \ddots & \\ & & 0 & 0 & & \sqrt{\omega_s}\end{array}\right]\quad (37)$$

and $\mathrm{Re}\,N$, $\mathrm{Im}\,N$ are the real and the imaginary parts of the matrix $N$, respectively. In particular, for one mode we obtain

$$\alpha=\hbar\begin{bmatrix}\omega^{-1}(N+\frac{1}{2}) & 0\\ 0 & \omega(N+\frac{1}{2})\end{bmatrix},\quad (38)$$

which corresponds to the characteristic function

$$\exp\left[-\frac{\hbar}{2}\left(N+\frac{1}{2}\right)(\omega^{-1}x^2+\omega y^2)\right].\quad (39)$$

## IV. THE ENTROPY AND THE CAPACITY

In this section we give a general formula for the von Neumann entropy of the arbitrary Gaussian density operator, and apply it to calculate the capacities for several concrete Gaussian channels.

### A. The entropy of the general Gaussian state

The following result can be found in [7], Proposition 2.1: *For arbitrary real symmetric matrix $\alpha$ there exists linear transformation $S:R\rightarrow SR=\tilde{R}$, such that $S\Delta S^T=\Delta$ and*

$$\tilde{\alpha}=S\alpha S^T=\left[\begin{array}{ccc|ccc}\alpha_1 & & 0 & 0 & & \\ & \ddots & & & \ddots & \\ 0 & & \alpha_s & & & 0\\ \hline 0 & & & \alpha_1 & & 0\\ & \ddots & & & \ddots & \\ & & 0 & 0 & & \alpha_s\end{array}\right]\quad (40)$$

The transformations satisfying $S\Delta S^T=\Delta$ preserve CCR and are called (linear) *canonical* transformations.

Let $\alpha$ be the correlation matrix of Gaussian density operator $\rho$ with $m=0$. It follows that

$$\operatorname{Tr}\rho\exp(i\tilde{R}^Tz)=\exp\left(-\frac{1}{2}z^T\tilde{\alpha}z\right)=\exp\left(-\frac{1}{2}\sum_{j=1}^{s}\alpha_j(x_j^2+y_j^2)\right).\quad (41)$$

The necessary and sufficient condition (30) is equivalent to $\alpha_j\geqslant\hbar/2,\ j=1,\ldots,s$. Therefore

$$\alpha_j=\hbar\left(N_j+\frac{1}{2}\right),\quad (42)$$

where $N_j\geqslant 0$. From Eq. (41)

$$\rho=\bigotimes_{j=1}^{s}\rho^{(j)},\quad (43)$$

where $\rho^{(j)}$ is the density operator for one degree of freedom, corresponding to the characteristic function (39) with $N=N_j,\ \omega=1$. From Sec. III A,

$$H(\rho)=\sum_{j=1}^{s} H(\rho^{(j)})=\sum_{j=1}^{s} g(N_j). \tag{44}$$

Consider the function $G$ which is defined by the following relation for $d\geqslant\frac{1}{2}$:

$$\begin{aligned}G(d^2)&=g\left(d-\frac{1}{2}\right)\\&=\left(d+\frac{1}{2}\right)\ln\left(d+\frac{1}{2}\right)-\left(d-\frac{1}{2}\right)\ln\left(d-\frac{1}{2}\right)\\&=\frac{1}{2}\ln d^2+\frac{1}{2}\ln\left(1-\frac{1}{4d^2}\right)+d\ln\frac{1+1/2d}{1-1/2d}\\&=\frac{1}{2}\ln d^2+\ln\left[1-\sum_{k=1}^{\infty}\frac{1}{2k(2k+1)}\left(\frac{1}{4d^2}\right)^k\right],\end{aligned} \tag{45}$$

and is a monotonously increasing concave function of $d^2$. Then

$$H(\rho)=\sum_{j=1}^{s} G\left(\left(\frac{\alpha_j}{\hbar}\right)^2\right)=\frac{1}{2}\mathrm{Sp}G(-(\Delta^{-1}\tilde{\alpha})^2). \tag{46}$$

Going back to the initial basis, we obtain

$$\Delta^{-1}\tilde{\alpha}=(S\Delta S^T)^{-1}S\tilde{\alpha}S^T=(S^T)^{-1}(\Delta^{-1}\alpha)S^T. \tag{47}$$

Hence $G(-(\Delta^{-1}\tilde{\alpha})^2)=(S^T)^{-1}G(-(\Delta^{-1}\alpha)^2)S^T$ and

$$H(\rho)=\frac{1}{2}\mathrm{Sp}\,G(-(\Delta^{-1}\alpha)^2), \tag{48}$$

which is the final expression for the entropy of the general Gaussian state. Note that it is valid for arbitrary skew-symmetric commutation matrix $\Delta$, not necessarily of the simplest canonical form (27). In order to aid understanding, we will show an example of how to use the formula.

*Example.* In the case of a general Gaussian density operator with one degree of freedom we have

$$\alpha=\begin{bmatrix}\alpha^{qq} & \alpha^{qp}\\ \alpha^{qp} & \alpha^{pp}\end{bmatrix}$$

with $\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2\geqslant\hbar^2/4$ [the last inequality is equivalent to Eq. (30)]. Then

$$-(\Delta^{-1}\alpha)^2=\frac{\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2}{\hbar^2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \tag{49}$$

and

$$H(\rho)=G\left(\frac{\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2}{\hbar^2}\right). \tag{50}$$

Geometrically, the argument of $G$ is equal to the squared area of the deviation ellipsoid

$$\pi\hbar z^T\alpha^{-1}z=1,\quad z=[x,y]^T,$$

for two-dimensional Gaussian distribution with the correlation matrix $\alpha$.

Note that in accordance with Eq. (32) the Gaussian state is pure if and only if $\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2=\hbar^2/4$. This comprises both coherent ($\alpha^{qq}=\hbar/2\omega$, $\alpha^{pp}=\hbar\omega/2$, $\alpha^{qp}=0$) and squeezed states in the real rather than complex parametrization. Let us show the relation between the real notation and the usual physical parametrization of squeezed states. Consider the squeezed state given by the vector $S(\zeta)|0\rangle$, where $\zeta=\gamma e^{i\theta}$ and $S(\zeta)=\exp\{[\zeta^* a^2-\zeta(a^\dagger)^2]/2\}$. Then $\alpha$ is represented by the squeezing parameters $\gamma,\theta$ as follows:

$$\begin{aligned}\alpha^{qq}&=\frac{\hbar}{2\omega}\{\cosh 2\gamma-\sinh 2\gamma\cos\theta\},\\ \alpha^{pp}&=\frac{\hbar\omega}{2}\{\cosh 2\gamma+\sinh 2\gamma\cos\theta\},\\ \alpha^{qp}&=\frac{\hbar}{2}\sinh 2\gamma\sin\theta.\end{aligned} \tag{51}$$

### B. Calculating the capacity

The capacity involves the maximization with respect to probability distribution of the signals. In this subsection we will show several examples of the maximization by using the formula given in the above subsection. Let $\rho(0)$ be a Gaussian density operator with $m=0$ and the correlation matrix $\alpha$. Let $P$ be a Gaussian probability distribution with correlation matrix $\beta$. Without loss of generality we may assume it has zero mean. The inequality (6) then takes the form (in terms of real variables)

$$\frac{1}{2}\sum_{j=1}^{s}(\omega_j^2\beta_{jj}^{qq}+\beta_{jj}^{pp})\leqslant E, \tag{52}$$

or

$$\mathrm{Sp}\,\varepsilon\beta\leqslant E, \tag{53}$$

where

$$\varepsilon=\left[\begin{array}{ccc|ccc}\frac{\omega_1^2}{2} & & 0 & 0 & & \\ & \ddots & & & \ddots & \\ 0 & & \frac{\omega_s^2}{2} & & & 0\\ \hline 0 & & & \frac{1}{2} & & 0\\ & \ddots & & & \ddots & \\ & & 0 & 0 & & \frac{1}{2}\end{array}\right]. \tag{54}$$

The mixture $\rho_P$ will again be a Gaussian density operator with zero mean and the correlation matrix $(\alpha+\beta)$ (see the Appendix). Thus the capacity of the general Gaussian channel is equal to

$$C=\max_{\beta\in B_1}\frac{1}{2}\mathrm{Sp}\,G(-[\Delta^{-1}(\alpha+\beta)]^2)-\frac{1}{2}\mathrm{Sp}\,G(-(\Delta^{-1}\alpha)^2) \tag{55}$$

where $B_1$ is the convex set of real positive matrices $\beta$, satisfying Eq. (53).

There are a few cases where the maximization in Eq. (55) can be made explicitly

(A) The case where $\rho(0)$ is the gauge-invariant density operator with diagonal matrix $N=\mathrm{diag}(N_j)$: see [4], where it is shown that the capacity is equal to the capacity of the corresponding quasiclassical "photonic channel." Namely, it is shown that

$$C=\max\sum_j g(N_j+m_j)-\sum_j g(N_j), \tag{56}$$

where the maximum is taken over all "mean photon numbers" $m_j\geqslant 0$, satisfying $\Sigma_j\hbar\omega_j m_j\leqslant E$. [In terms of the matrix $\beta$ the quantities $m_j=(1/2\hbar\omega_j)(\omega_j^2\beta_{jj}^{qq}+\beta_{jj}^{pp})$]. The problem (56) has the solution

$$m_j=(N_j(\lambda)-N_j)_+, \tag{57}$$

where $N_j(\lambda)=(\exp\lambda\hbar\omega_j-1)^{-1}$,$(x)_+$ denotes $x$ if $x\geqslant 0$, and 0 otherwise, and $\lambda$ is chosen in such a way that $\Sigma_j\hbar\omega_j m_j=E$ [8,4]. This is a modification of the classical "water-filling solution" [9] with the quantum "water level" given by the one-dimensional Planck distribution $N_j(\lambda)$. In this case the matrix $N$ commutes with $\varepsilon$ and $\Delta$, making the calculation possible.

In particular, for one mode

$$C=g(N+N_s)-g(N), \tag{58}$$

where $N=\mathrm{Tr}\,\rho(0)a^\dagger a$, $N_s=E/\hbar\omega$. Thus the quantity, which was conjectured long ago as the upper bound for the capacity of the Gaussian channel [10], is in fact the capacity.

(B) The case of general $\rho(0)$ in one mode, in particular, the case of a squeezed state.

(B1) First we shall calculate the capacity (55) for the given correlation matrix $\alpha$ corresponding to $\rho(0)$. When we use the Eq. (50), Eq. (55) can be rewritten as follows:

$$C=\max_{\beta\in B_1}G\left(\frac{(\alpha^{qq}+\beta^{qq})(\alpha^{pp}+\beta^{pp})-(\alpha^{qp}+\beta^{qp})^2}{\hbar^2}\right)-G\left(\frac{(\alpha^{qq})(\alpha^{pp})-(\alpha^{qp})^2}{\hbar^2}\right), \tag{59}$$

where $B_1$ is a set of real positive $2\times 2$ matrices $\beta$ satisfying the condition (53), i.e.,

$$\frac{1}{2}(\omega^2\beta^{qq}+\beta^{pp})\leqslant E, \tag{60}$$

$$\beta^{qq}\beta^{pp}\geqslant(\beta^{qp})^2. \tag{61}$$

Because $G(d^2)$ is a monotonously increasing function of $d^2$, the capacity is given by

$$C=G(d_{\max}^2)-G\left(\frac{\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2}{\hbar^2}\right), \tag{62}$$

where

$$d_{\max}^2=\frac{1}{\hbar^2}\max_{\beta\in B_1}[(\alpha^{qq}+\beta^{qq})(\alpha^{pp}+\beta^{pp})-(\alpha^{qp}+\beta^{qp})^2], \tag{63}$$

which can be calculated in the following way.

We can replace Eq. (60) by the equation

$$\frac{1}{2}(\omega^2\beta^{qq}+\beta^{pp})=E \tag{64}$$

without loss of generality. When we use Eqs. (61) and (64), the variable region of $\beta^{qq}$ and $\beta^{pp}$ can be determined by the following inequality, which represents an elliptic region:

$$\beta^{qq}\left(\beta^{qq}-\frac{2E}{\omega^2}\right)+\left(\frac{\beta^{qp}}{\omega}\right)^2\leqslant 0. \tag{65}$$

Thus $\beta^{qq}$ and $\beta^{qp}$ can be represented with the parameters $r$ and $\phi$ as follows:

$$\beta^{qq}=r\cos\phi+\frac{E}{\omega^2},\quad \beta^{qp}=r\omega\sin\phi, \tag{66}$$

$$0\leqslant r\leqslant\frac{E}{\omega^2},\quad 0\leqslant\phi<2\pi. \tag{67}$$

Substituting Eq. (66) into Eq. (63), we have

$$d_{\max}^2=\frac{1}{\omega^2\hbar^2}\left(E+\frac{1}{2}(\omega^2\alpha^{qq}+\alpha^{pp})\right)^2-\frac{1}{\hbar^2}\min_{(r,\phi)\in\tilde{B}_1}\left\{\left[r\omega\cos\phi+\frac{1}{2}\left(\omega\alpha^{qq}-\frac{\alpha^{pp}}{\omega}\right)\right]^2+(r\omega\sin\phi+\alpha^{qp})^2\right\}, \tag{68}$$

where $\tilde{B}_1$ is the set of parameters $(r,\phi)$ satisfying Eq. (67). Now let $D$ be the disk with the radius $E/\omega$ and the center 0. Then the second term of Eq. (68) is equal to the distance between the point $(-\frac{1}{2}(\omega\alpha^{qq}-\alpha^{pp}/\omega),-\alpha^{qp})$ and the disk $D$, and hence it vanishes if and only if

$$\left[\frac{1}{2}\left(\omega\alpha^{qq}-\frac{\alpha^{pp}}{\omega}\right)\right]^2+(\alpha^{qp})^2\leqslant\left(\frac{E}{\omega}\right)^2. \tag{69}$$

If Eq. (69) holds, then the capacity is given by

$$C=G\left(\frac{1}{\hbar^2\omega^2}\left(E+\frac{1}{2}(\omega^2\alpha^{qq}+\alpha^{pp})\right)^2\right)-G\left(\frac{\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2}{\hbar^2}\right), \tag{70}$$

or

$$C=g(N+N_s)-H(\rho(0)),\qquad(71)$$

where $N_s=E/\hbar\omega$, $N=\mathrm{Tr}\,\rho(0)a^{\dagger}a=(\omega^2\alpha^{pp}+\alpha^{qq})/2\hbar\omega-\frac{1}{2}$, and $H(\rho(0))$ is given by Eq. (50).

In particular, if $\rho(0)$ is pure, that is a squeezed state [i.e., $\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2=\hbar^2/4$], with parameters satisfying Eq. (69), the capacity becomes

$$G\left(\frac{1}{\hbar^2\omega^2}\left(E+\frac{1}{2}(\omega^2\alpha^{qq}+\alpha^{pp})\right)^2\right),\qquad(72)$$

or

$$C=g(N+N_s).\qquad(73)$$

On the other hand, if Eq. (69) does not hold, the second term of Eq. (68) becomes

$$\frac{1}{\hbar^2}\left\{\sqrt{\left[\frac{1}{2}\left(\omega\alpha^{qq}-\frac{\alpha^{pp}}{\omega}\right)\right]^2+(\alpha^{qp})^2}-\frac{E}{\omega}\right\}^2.\qquad(74)$$

Then the capacity is given by

$$C=G\left(\frac{1}{\hbar^2\omega^2}\left\{\left(E+\frac{1}{2}(\omega^2\alpha^{qq}+\alpha^{pp})\right)^2-\left[\sqrt{\left(\frac{1}{2}(\omega^2\alpha^{qq}-\alpha^{pp})\right)^2+\omega^2(\alpha^{qp})^2}-E\right]^2\right\}\right)-G\left(\frac{\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2}{\hbar^2}\right).\qquad(75)$$

In particular, if $\rho(0)$ is squeezed state, by substituting $(\alpha^{qp})^2=\alpha^{qq}\alpha^{pp}-\hbar^2/4$ in Eq. (75), we have

$$C=G\left(\frac{1}{\hbar^2\omega^2}\left(E(\omega^2\alpha^{qq}+\alpha^{pp})+E\sqrt{(\omega^2\alpha^{qq}+\alpha^{pp})^2-\hbar^2\omega^2}+\frac{\hbar^2\omega^2}{4}\right)\right)=G\left(N_s(2N+1+2\sqrt{N^2+N})+\frac{1}{4}\right).\qquad(76)$$

Now let us consider the capacity $C$ as the function of the noise correlation matrix $\alpha$. Then it is easy to see that $C$ has a minimum when $\alpha^{qq}=\hbar/2\omega$, $\alpha^{pp}=\hbar\omega/2$, $\alpha^{qp}=0$ [i.e., $\rho(0)$ is coherent state]. In other words, using squeezing states under constrained input energy does increase the capacity. The situation where the capacity is given by different expressions depending on whether or not the noise parameters fulfill a certain inequality may be interpreted as a further noncommutative generalization of the ''quantum water filling.'' The graph of the capacity as the function of squeezing parameter $\gamma$ for $\theta=0$ and for $N_s=1$ is presented in Fig. 1.

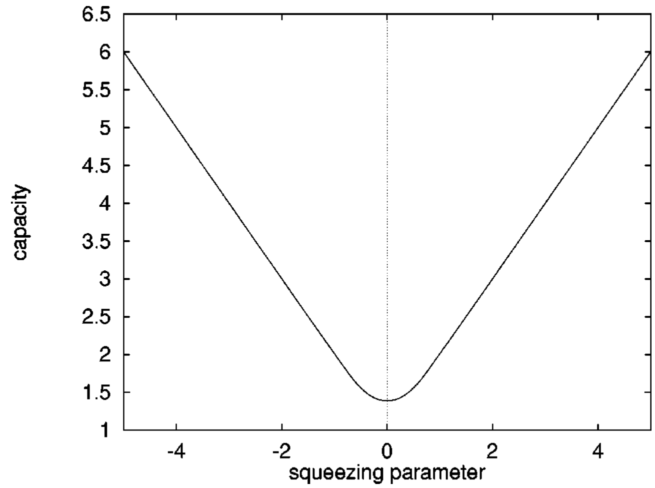


FIG. 1. Dependence of the capacity of the squeezed state channel with constrained signal power, $N_s=1$, on the squeezing parameter $\gamma$.

(B2): In (B1) we calculated the capacity under the *input* power constraint (60). Here we shall calculate the capacity under the *output* (signal plus noise) power constraint:

$$\tfrac{1}{2}[\omega^2(\alpha^{qq}+\beta^{qq})+\beta^{pp}+\alpha^{pp}]\leqslant\hbar\omega(N_t+\tfrac{1}{2}),\qquad(77)$$

where $N_t$ is the total mean photon number in the output. In this case we can obtain the capacity in the same way as (B1), by replacing the energy bound $E$ with

$$\omega\hbar(N_t+\tfrac{1}{2})-\tfrac{1}{2}(\omega^2\alpha^{qq}+\alpha^{pp}),\qquad(78)$$

provided $\alpha$ is such that this quantity is positive. Then, Eq. (69) is replaced by

$$\left[\frac{1}{2}\left(\omega\alpha^{qq}-\frac{\alpha^{pp}}{\omega}\right)\right]^2+(\alpha^{qp})^2\leqslant\left\{\frac{1}{\omega}\left[\omega\hbar\left(N_t+\frac{1}{2}\right)-\frac{1}{2}(\omega^2\alpha^{qq}+\alpha^{pp})\right]\right\}^2.\qquad(79)$$

Thus, if Eq. (79) holds, the capacity is given by

$$C=G\left(\left(N_t+\frac{1}{2}\right)^2\right)-G\left(\frac{(\alpha^{qq})(\alpha^{pp})-(\alpha^{qp})^2}{\hbar^2}\right).$$

In particular, if $\rho(0)$ is pure [i.e., $\alpha^{qq}\alpha^{pp}-(\alpha^{qp})^2=\hbar^2/4$] and satisfies Eq. (79), we have

$$C=G((N_t+\tfrac{1}{2})^2)=g(N_t),\qquad(80)$$

which is equal to the Yuen-Ozawa bound (von Neumann entropy of the thermal state). In accordance with [11], in this case use of squeezed states cannot improve capacity. It is interesting that this does not become worse, provided the condition (79) holds.

Next we shall represent the condition (79) by a squeezing parameter. Let $\rho(0)$ be a squeezed state with the vector $S(\zeta)|0\rangle$, where $\zeta=\gamma e^{i\theta}$. Let us consider the case where $\theta=0$. Then from Eq. (51) we have

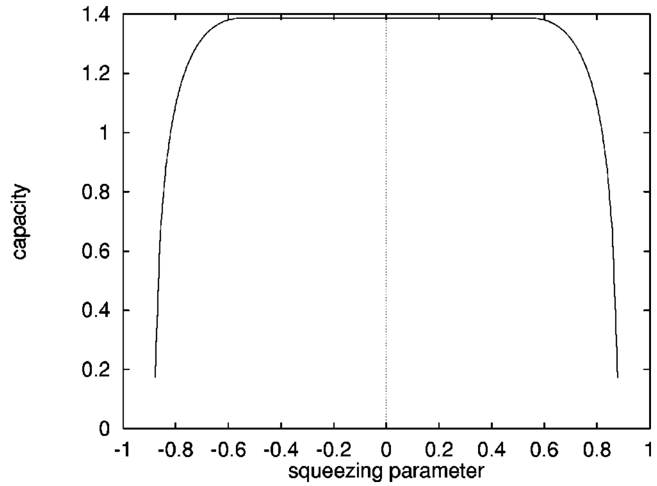


FIG. 2. Dependence of the capacity of the complex amplitude squeezed state channel with constrained output power, $N_t = 1$, on the squeezing parameter $\gamma$.

$$\alpha^{qq} = \frac{\hbar}{2\omega} e^{-2\gamma}, \quad \alpha^{pp} = \frac{\hbar\omega}{2} e^{2\gamma}, \quad \alpha^{qp} = 0. \tag{81}$$

Substituting Eq. (81) into Eq. (79), we get the following condition:

$$\ln\sqrt{\frac{1}{2N_t+1}} \leq \gamma \leq \ln\sqrt{2N_t+1}. \tag{82}$$

Thus we obtain the same capacity as for number or coherent states by using squeezed states with the squeezing parameter $\zeta = \gamma$ satisfying Eq. (82). Figure 2 illustrates capacity $C$ as a function of squeezing parameter in the case of $N_t = 1$.

Furthermore, we shall consider whether there is a value of the squeezing parameter which makes the capacity equal to the Yuen-Ozawa bound or not, when the coherent amplitude of the squeezing state is restricted to a real number, i.e.,

$$\beta^{pp} = 0, \quad \beta^{qp} = 0. \tag{83}$$

Now we can suppose that the equality of Eq. (77) holds without loss of generality. Then by substituting Eq. (83) into Eq. (77), we have

$$\tfrac{1}{2}[\omega^2(\alpha^{qq} + \beta^{qq}) + \alpha^{qq}] = \hbar\omega(N_t + \tfrac{1}{2}). \tag{84}$$

So for the given $\alpha$, the correlation matrix $\beta$ has the entries

$$\beta_0^{pp} = 0, \quad \beta_0^{qp} = 0,$$

$$\beta_0^{qq} = \frac{2\hbar}{\omega}\left(N_t + \frac{1}{2}\right) - \frac{1}{\omega^2}(\omega^2\alpha^{qq} + \alpha^{pp}), \tag{85}$$

provided $\beta_0^{qq} > 0$. Therefore Eq. (59) can be rewritten simply as follows:

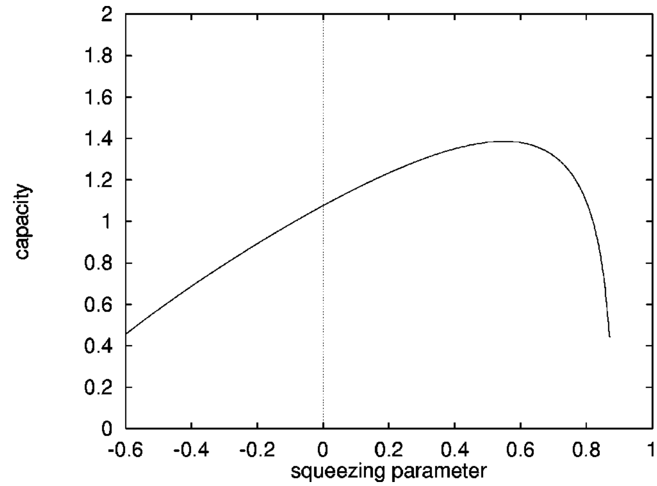


FIG. 3. Dependence of the capacity of the real amplitude squeezed state channel with constrained output power, $N_t = 1$, on the squeezing parameter $\gamma$.

$$C = G\left(\frac{(\alpha^{qq} + \beta_0^{qq})(\alpha^{pp} + \beta_0^{pp}) - (\alpha^{qp} + \beta_0^{qp})^2}{\hbar^2}\right) - G\left(\frac{(\alpha^{qq})(\alpha^{pp}) - (\alpha^{qp})^2}{\hbar^2}\right). \tag{86}$$

Substituting Eqs. (81) and (85) into Eq. (86), we have

$$C = G(-[\tfrac{1}{2}e^{2\gamma} - (N_t + \tfrac{1}{2})]^2 + (N_t + \tfrac{1}{2})^2). \tag{87}$$

Thus we find that the capacity $C$ takes the maximum value, when the squeezing parameter $\gamma$ is equal to

$$\gamma_0 = \ln\sqrt{2N_t+1}. \tag{88}$$

When $\rho(0)$ is the squeezed state with the parameter (88) and the receiver is homodyne, we have

$$I_{\max} = \ln(1 + 2N_t), \tag{89}$$

which is the maximum value for any squeezing parameter. So $I_{\max}$ is greater than that in the case of a coherent state channel. But it is less than the Yuen-Ozawa bound. We have seen that the capacity $C$ becomes equal to the Yuen-Ozawa bound if and only if Eq. (88) holds. Thus use of squeezed states can improve the maximum mutual information under the output power constraint, but cannot improve the capacity. Figure 3 illustrates the capacity $C$ as a function of squeezing parameter in the case of $N_t = 1$.

Note that the value of the squeezing parameter $\gamma_0$ is equal to that for the maximum signal-to-noise ratio: $4N_t(N_t+1)$ [12].

## V. CONCLUSION

In this paper we considered a memoryless communication channel representing classical signal plus background quantum Gaussian noise. We described a broad class of quantum Gaussian states, which comprises both coherent and squeezed states, as well as their Gaussian mixtures, and gave an expression for von Neumann's entropy of a general

Gaussian state. Basing on this formula and on the recently obtained formula for the classical capacity of a quantum channel, we calculated the capacity of the squeezed state channel, showing quantitatively the information-theoretic advantage over the channel using coherent states, under constrained input signal power. The capacity is achieved by a Gaussian input distribution satisfying a certain condition which can be interpreted as a quantum counterpart of the classical ''water-filling condition.''

On the contrary, using squeezed states under constrained output power cannot increase the capacity, in accordance with the result of [11], although somewhat surprisingly, it does not make it worse, provided the squeezing is not too large.

*Note added in proof.* The formula (50) for the entropy of a Gaussian state with one degree of freedom agrees with one obtained in G. S. Agarwal, Phys. Rev. A **3**, 828 (1971). The authors are grateful to Professor G. S. Agarwal for bringing this reference to their attention.


## ACKNOWLEDGMENTS

A.S.H. acknowledges the hospitality of the Research Center of Tamagawa University, where a substantial part of this paper was prepared.


## APPENDIX

*Proof of the property* (i). Let $\rho$ be a density operator with finite second moments. According to the property (iv) there is a Gaussian density operator $\tilde{\rho}$ with the same first and second moments. We then have

$$H(\tilde{\rho})-H(\rho)=\mathrm{Tr}\,\rho(\ln\rho-\ln\tilde{\rho})+\mathrm{Tr}(\rho-\tilde{\rho})\ln\tilde{\rho}. \quad \text{(A1)}$$

The first term on the right is relative entropy, which is known to be nonnegative. The second term is zero, because $\tilde{\rho}$ and $\rho$ have the same first and second moments, and $\ln\tilde{\rho}$ is a polynomial of the second order in $a, a^\dagger$. This follows from the fact that after certain linear canonical transformation of $a, a^\dagger$, the density operator $\tilde{\rho}$ can be represented as the tensor product of elementary operators (16), which can be put in the form

$$\rho_j(0)=\frac{1}{N_j+1}\left(\frac{N_j}{N_j+1}\right)^{a_j^\dagger a_j}, \quad \text{(A2)}$$

making obvious that $\ln\rho_j(0)$ is a second order polynomial in $a, a^\dagger$. Thus $H(\tilde{\rho})-H(\rho)\geqslant 0$.

*Proof of the property* (ii). Let $P$ be a Gaussian probability distribution with zero mean and covariance matrix $\beta$ for the variables $m_1^q,\dots,m_s^q,m_1^p,\dots,m_s^p$. The quantum characteristic function of the density operator $\rho_P$ is

$$\begin{aligned}\mathrm{Tr}\,\rho_P V(z)&=\int \exp(im^T z-\tfrac{1}{2}z^T\alpha z)P(dm)\\&=\varphi_P(z)\exp(-\tfrac{1}{2}z^T\alpha z),\end{aligned} \quad \text{(A3)}$$

where $\varphi_P(z)=\exp(-\frac{1}{2}z^T\beta z)$ is the classical characteristic function of the probability distribution $P$. This proves (ii).

*Proof of the property* (iii). The state with the density operator $\rho$ is pure if and only if $\mathrm{Tr}\,\rho^2=1$. However, by the quantum Parceval identity (see [7]), for Gaussian $\rho$

$$\begin{aligned}\mathrm{Tr}\,\rho^2&=(2\pi)^{-s}\int |\mathrm{Tr}\,\rho V(z)|^2 d^{2s}z\\&=(2\pi\hbar)^{-s}\int\cdots\int \exp(-z^T\alpha z)dx_1\cdots dx_s dy_1\cdots dy_s\\&=\sqrt{\frac{\det(2\alpha)}{\hbar^{2s}}}.\end{aligned} \quad \text{(A4)}$$

Thus $\mathrm{Tr}\,\rho^2=1$ if and only if Eq. (33) holds.


[1] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. Wootters, Phys. Rev. A **54**, 1869 (1996).

[2] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 296 (1998).

[3] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).

[4] A. S. Holevo, Tamagawa University Research Review **4**, 1 (1998); see also LANL e-print quant-ph/9809023.

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

[6] A. S. Holevo, IEEE Trans. Inf. Theory **IT21**, 533 (1975).

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

[8] D. S. Lebedev and L. B. Levitin, Inf. Control **9**, 1 (1966).

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

[10] J. P. Gordon, in *Quantum Electronics and Coherent Light*, Proceedings of the International School of Physics ''Enrico Fermi,'' Course XXXI, edited by P. A. Miles (Academic Press, New York, 1964), pp. 156–181.

[11] H. P. Yuen and M. Ozawa, Phys. Rev. Lett. **70**, 363 (1993).

[12] H. P. Yuen, Phys. Lett. **56A**, 105 (1976).