Information capacity of continuous variable measurement channel

A S Holevo[1] and A A Kuznetsova

Steklov Mathematical Institute, Gubkina 8, 119991 Moscow, Russia

E-mail: holevo@mi-ras.ru and kuznetsova.a.a@bk.ru

Abstract

The present paper is devoted to investigation of the classical capacity of infinite-dimensional continuous variable quantum measurement channels. A number of usable conditions are introduced that enable us to apply previously obtained general results to specific models, in particular, to the multi-mode bosonic Gaussian measurement channels. An explicit formula for the classical capacity of the Gaussian measurement channel is obtained in this paper without assuming the global gauge symmetry, solely under certain 'threshold condition'. The result is illustrated by the capacity computation for a one-mode squeezed-noise heterodyne measurement channel.

Keywords: quantum information theory, quantum measurement channel, classical capacity, continuous variable system, Gaussian observable, threshold condition

1. Introduction

From the viewpoint of information theory measurements are peculiar communication channels that transform input quantum states into classical output data. As such, they are described by the information capacity which is the most important quantity characterizing their ultimate information-processing performance. The present work develops investigation of the capacities of quantum measurement channels, initiated in [1–4]. The emphasis here is on the measurements with continuous multi-dimensional output. It is well known (see, e.g., [3]) that channels with continuous classical output (in contrast to discrete output) in principle cannot be extended to quantum channels (i.e. maps with the quantum input and output). This prevents from a direct use of the well developed quantum Shannon theory for channels and thus requires a separate study. In particular, this remark fully applies to bosonic Gaussian measurement channels, which are of great both theoretical and practical interest.

In [5] a proof of the coding theorem for the classical capacity of a measurement channel with arbitrary output alphabet was given in the most general setting. In the present work

[1] Author to whom any correspondence should be addressed.

a number of applicable conditions are developed that make it possible to implement the general results, in particular, to bring the capacity calculations to the quite specific formulas. In proposition 1 of section 2 a convenient expression is obtained for the energy-constrained classical capacity of a measurement channel in terms of the differential entropy. This expression is further specified in proposition 2 for irreducibly covariant measurement channels. Section 3 is devoted to Gaussian measurement channels. Theorem 1 is proved, giving an explicit formula for the energy-constrained classical capacity of the Gaussian measurement channel without assuming the global gauge symmetry, solely under the 'threshold condition' (21). This result generalizes theorem 1 from [6], in which the globally gauge-covariant case was considered. The result is illustrated by an example of the single-mode squeezed-noise heterodyne measurement.

2. Capacity of the measurement channel and the output differential entropy

Let $\mathcal{H}$ be a separable Hilbert space. By $\mathfrak{B}(\mathcal{H})$ we denote the algebra of all bounded operators in $\mathcal{H}$, $\mathfrak{T}(\mathcal{H})$ is the Banach space of trace-class operators, $\mathfrak{S}(\mathcal{H})$ is the convex subset of *density operators*, i.e. positive operators with unit trace, also called *quantum states*.

We also introduce a measure space $(\Omega, \mathcal{F}, \mu)$, where Ω is a complete separable metric space, $\mathcal{F}$ is a σ-algebra of its subsets, μ is a σ-finite measure on $\mathcal{F}$.

Definition 1. Probability operator-valued measure (POVM) on Ω is a family $M = \{M(A), A \in \mathcal{F}\}$ of bounded Hermitian operators in $\mathcal{H}$, satisfying the following conditions:

(a) $M(A) \geqslant 0; A \in \mathcal{F}$;
(b) $M(\Omega) = I$, where I is the unit operator in $\mathcal{H}$;
(c) for arbitrary countable decomposition $A = \cup A_i$, $(A_i \in \mathcal{F}, A_i \cap A_j = \emptyset, i \neq j)$ the relation $M(A) = \sum_i M(A_i)$ holds in the sense of weak convergence of operators.

The POVM defines a *quantum observable* with values in Ω. *The probability distribution* of the observable M in the state ρ is given by the formula

$$P_\rho(A) = \mathrm{Tr}\, \rho M(A), \quad A \in \mathcal{F}. \tag{1}$$

For brevity, we sometimes write $P_\rho(\mathrm{d}\omega) = \mathrm{Tr}\, \rho M(\mathrm{d}\omega)$. As it is known, there is a unique POVM, for which P_ρ is given by the formula (1).

Definition 2. *Measurement channel* $\mathcal{M}$ is an affine map $\rho \to P_\rho(\mathrm{d}\omega)$ of the convex set of quantum states $\mathfrak{S}(\mathcal{H})$ into the set of probability distributions on Ω.

The purpose of this work is to study the information characteristics of the channel $\mathcal{M}$, in particular, its capacity for transmitting the classical information, with a natural energy restriction at the input, basing on general expressions obtained previously in [5, 7].

Lemma 1. [5] *For arbitrary observable $M(\mathrm{d}\omega)$ with values in Ω there exists a σ-finite measure μ on Ω, such that for any density operator $\rho \in \mathfrak{S}(\mathcal{H})$ the probability distribution $P_\rho(\mathrm{d}\omega) = \mathrm{Tr}\, \rho M(\mathrm{d}\omega)$ has a density $p_\rho(\omega)$ with respect to measure μ.*

Therefore the measurement channel can be considered as an affine map $\mathcal{M} : \rho \to p_\rho(\omega)$, and we will write $p_\rho = \mathcal{M}(\rho)$.

Definition 3. A Borel probability measure π on $\mathfrak{S}(\mathcal{H})$ will be called *ensemble*. The *average state* of the ensemble π, determined by the Bochner integral in $\mathfrak{T}(\mathcal{H})$ [8]

$$\bar{\rho}_\pi = \int\limits_{\mathfrak{S}(\mathcal{H})} \rho\pi(\mathrm{d}\rho), \tag{2}$$

is just the barycenter of the measure π.

Note that in specific applications the measure π is usually concentrated on some parametrically given subset of states. In these cases, it is more convenient to define the ensemble as a pair $\{\pi(\mathrm{d}x), \rho_x\}$, where the parameter x runs over a complete separable metric space X, $\pi(\mathrm{d}x)$ is a probability measure on X, $\{\rho_x, x \in X\}$ is a measurable family of states.

For a given ensemble π and a measurement channel $\mathcal{M}$ the Shannon mutual information between the input and the output of the channel can be defined by the formula

$$I(\pi, \mathcal{M}) = \int_{\mathfrak{S}(\mathcal{H})} h(\mathcal{M}(\rho)\|\mathcal{M}(\bar{\rho}_\pi))\pi(\mathrm{d}\rho),$$

where

$$h(\mathcal{M}(\rho)\|\mathcal{M}(\bar{\rho}_\pi)) = \int_\Omega p_\rho(\omega) \log \frac{p_\rho(\omega)}{p_{\bar{\rho}_\pi}(\omega)} \mu(\mathrm{d}\omega)$$

is the classical relative entropy. The functional $I(\pi, \mathcal{M})$ is well defined and takes values in $[0; +\infty]$.

We introduce the generalized differential entropy of a probability density $p(\omega)$ on (Ω, μ) by the relation (see [9], where μ is the Lebesgue measure on $\mathbb{R}^s$)

$$h(p) = -\int_\Omega p(\omega) \log p(\omega)\mu(\mathrm{d}\omega),$$

provided the integral exists, including the values $\pm\infty$. If $|h(\mathcal{M}(\bar{\rho}_\pi))| < \infty$ then

$$I(\pi, \mathcal{M}) = \int_{\mathfrak{S}(\mathcal{H})}\int_\Omega p_\rho(\omega) \left(\log p_\rho(\omega) - \log p_{\bar{\rho}_\pi}(\omega)\right) \mu(\mathrm{d}\omega)\pi(\mathrm{d}\rho)$$

$$= h(\mathcal{M}(\bar{\rho}_\pi)) - \int_{\mathfrak{S}(\mathcal{H})} h(\mathcal{M}(\rho))\pi(\mathrm{d}\rho),$$

where $h(\mathcal{M}(\rho)) = h(p_\rho)$ is the differential entropy of the channel $\mathcal{M}$ output probability density.

In the case of infinite-dimensional $\mathcal{H}$ one usually introduces a constraint onto the input states of the channel (otherwise the capacity may be infinite). Let H be a positive self-adjoint (in general unbounded) operator in the space $\mathcal{H}$, with the spectral decomposition $H = \int_0^\infty \lambda\mathrm{d}P(\lambda)$, where $P(\lambda)$ is the spectral function. In specific applications, the role of H is played by the energy operator of a quantum system at the input of the channel. We introduce the subset of states

$$\mathfrak{S}_E = \{\rho : \mathrm{Tr}\,\rho H \leqslant E\}, \tag{3}$$

where E is a positive constant, and the trace is understood as the integral (for more detail see [10])

$$\mathrm{Tr}\,\rho H = \int_0^\infty \lambda\mathrm{d}(\mathrm{Tr}\,\rho P(\lambda)).$$

Consider the energy constraint on the input ensemble π defined by the relation $\mathrm{Tr}\ \bar{\rho}_\pi H \leqslant E$. According to theorem 1 from [5], the classical capacity of the measurement channel $\mathcal{M}$ with the input constraint is given by the relation

$$C(\mathcal{M}, H, E) = \sup_{\pi':\mathrm{Tr}\ \bar{\rho}_{\pi'}H \leqslant E} I(\pi', \mathcal{M}),$$

where $\pi' = \{\pi'_x, \rho_x'\}$ runs through the finite ensembles. Then, obviously,

$$C(\mathcal{M}, H, E) \leqslant \sup_{\pi:\mathrm{Tr}\ \bar{\rho}_\pi H \leqslant E} I(\pi, \mathcal{M}), \tag{4}$$

where the supremum is taken over all ensembles.

Proposition 1. *Let $\mathcal{M} : \rho \to p_\rho(\omega)$ be a measurement channel such that for any ρ the density $p_\rho(\omega)$ is bounded.*

Let a nonnegative function $H_\mathrm{c}(\omega)$ be given on Ω that satisfies the conditions

$$\int_\Omega \mathrm{e}^{-\theta H_\mathrm{c}(\omega)} \mu(\mathrm{d}\omega) < \infty, \tag{5}$$

for some $\theta > 0$, and

$$\int_\Omega H_\mathrm{c}(\omega) p_\rho(\omega) \mu(\mathrm{d}\omega) < \infty, \quad \forall \rho \in \mathfrak{S}_E, \tag{6}$$

then

$$h(p_\rho) < \infty, \quad \forall \rho \in \mathfrak{S}_E. \tag{7}$$

In this case, the capacity $C(\mathcal{M}, H, E)$ is given by the formula

$$C(\mathcal{M}, H, E) = \sup_{\pi:\mathrm{Tr}\ \bar{\rho}_\pi H \leqslant E} \left[h(\mathcal{M}(\bar{\rho}_\pi)) - \int_{\mathfrak{S}(\mathcal{H})} h(\mathcal{M}(\rho)) \pi(\mathrm{d}\rho) \right], \tag{8}$$

where the supremum is taken over all ensembles.

Proof. It is shown in [6] that the value $-\infty$ is excluded for the output differential entropy due to the fact that the density p_ρ is bounded. We show that under the condition (6) the value $+\infty$ is also excluded.

On the output space Ω, we consider the probability distribution with the density

$$p_\mathrm{H}(\omega) = m^{-1} \mathrm{e}^{-\theta H_\mathrm{c}(\omega)}, \quad m = \int_\Omega \mathrm{e}^{-\theta H_\mathrm{c}(\omega)} \mu(\mathrm{d}\omega) \tag{9}$$

with respect to the measure μ, then

$$h(p_\rho) = -h(p_\rho \| p_\mathrm{H}) + \theta \int_\Omega p_\rho(\omega) H_\mathrm{c}(\omega) \mu(\mathrm{d}\omega) + \log m. \tag{10}$$

This implies (7) due to the non-negativity of the relative entropy $h(p_\rho \| p_\mathrm{H})$ and the inequality (6).

Consider a finite decomposition $\mathcal{V} = \{V_i\}$ of the output space Ω and the measurement channel $\mathcal{M}_\mathcal{V}$ described by POVM $\{M(V_i)\}$, which is embedded into a quantum channel with finite-dimensional output.

For the capacity of $\mathcal{M}_\mathcal{V}$ we have

$$C(\mathcal{M}, H, E) \geqslant C(\mathcal{M}_\mathcal{V}, H, E) = \sup_{\pi: \mathrm{Tr}\, \bar{\rho}_\pi H \leqslant E} I(\pi, \mathcal{M}_\mathcal{V}),$$

where in the last equality we take a supremum over all ensembles and use the fact that the measurement channel $\mathcal{M}_\mathcal{V}$ can be considered as a quantum entanglement-breaking channel [7]. By taking supremum over the decompositions $\mathcal{V}$ and using the fact that $\sup_\mathcal{V} I(\pi, \mathcal{M}_\mathcal{V}) = I(\pi, \mathcal{M})$ (theorem 2.2 in [11]), we obtain the lower estimate which, together with the upper estimate (4) gives (8). □

Definition 4. Let G be a locally compact group, acting continuously on the transitive G-space Ω, let also $V : g \to V_g$ be a continuous (projective) unitary representation of the group G in the Hilbert space $\mathcal{H}$. A POVM $M = \{M(A), A \in \mathcal{F}\}$ in $\mathcal{H}$ is called *covariant* under V, if

$$V_g^* M(A) V_g = M(g^{-1}A), \quad A \in \mathcal{F}.$$

The following statement holds for the corresponding covariant measurement channel (an analogue of proposition 1 from [12]). We assume that the channel $\mathcal{M}$ satisfies the conditions of proposition 1 so that the formula (8) holds.

Proposition 2. *Let the following conditions be satisfied for the measurement channel $\mathcal{M}$ corresponding to the observable M which is covariant under an irreducible square integrable representation $g \to V_g$ of an unimodular group G:*

(a) $\displaystyle\sup_{\rho: \mathrm{Tr}\, \rho H \leqslant E} h(\mathcal{M}(\rho))$ *is attained on a state ρ_E^0;*
(b) $\displaystyle\inf_\rho h(\mathcal{M}(\rho))$ *is attained on a state ρ_0;*
(c) there exists a Borel probability measure π_E^0 on G such that

$$\rho_E^0 = \int_G V_g \rho_0 V_g^* \pi_E^0(\mathrm{dg}).$$

Then the capacity of the channel $\mathcal{M}$ is given by the formula

$$C(\mathcal{M}, H, E) = h(\mathcal{M}(\rho_E^0)) - h(\mathcal{M}(\rho_0)), \tag{11}$$

and it is attained on the ensemble $\{\pi_E^0(\mathrm{dg}), V_g \rho_0 V_g^*\}$.

Proof. From (8)

$$C(\mathcal{M}, H, E) \leqslant h(\mathcal{M}(\rho_E^0)) - h(\mathcal{M}(\rho_0)). \tag{12}$$

For a fixed point $\omega_0 \in \Omega$, consider its stationary subgroup G_0. According to theorem 4.8.3 of [13] the relation

$$M(A) = \int_A V_{g'} P_0 V_{g'}^* \mu(\mathrm{d}\omega), \quad \text{where } \omega = g'\omega_0 \tag{13}$$

establishes one-to-one correspondence between the measurements $\{M(A)\}$, covariant with respect to the irreducible square integrable representation $g \to V_g$ of the unimodular group G, and the density operators P_0, commuting with operators $\{V_g, g \in G_0\}$. Here $\mu(\mathrm{d}\omega)$ is properly normalized G-invariant measure on Ω, and the integral (13) is understood in the weak sense.

Using the representation (13), the covariance of the measurement channel $\mathcal{M}$ and the invariance of the measure μ, we get

$$h(\mathcal{M}(V_g \rho_0 V_g^*)) = h(\mathcal{M}(\rho_0)),$$

for any $g \in G$. Substituting the ensemble $\{\pi_E^0(\mathrm{d}g), V_g \rho_0 V_g^*\}$ in the formula (8), we obtain that the upper bound (12) is achieved, which completes the proof. □

3. The classical capacity of general Gaussian observable

The measurement channels we consider in this section correspond to general Gaussian observables in the sense of [14] (linear measurements in [13], see also [6] for the gauge covariant case). In what follows the quantum input space $\mathcal{H}$ will be the Hilbert space of an irreducible representation of the canonical commutation relations (A2) (see appendix), and the classical output space $\Omega = Z = \mathbb{R}^{2s}$.

We will consider the general Gaussian measurement channel $\rho \to \widetilde{\mathcal{M}}[\rho]$, described by POVM on Z

$$\widetilde{M}(\mathrm{d}^{2s}z) = D(Kz)\rho_\beta D(Kz)^* \frac{|\det K|\, \mathrm{d}^{2s}z}{(2\pi)^s}; \quad z \in \mathbb{R}^{2s}, \tag{14}$$

where z is $2s$-dimensional real vector running in the symplectic space $(\mathbb{R}^{2s}, \Delta)$, $D(z) = W(-\Delta^{-1}z)$ are the displacement operators (see appendix) satisfying the equation that follows from the canonical commutation relations (A2)

$$D(z)^* W(w) D(z) = \exp\left(\mathrm{i}w^t z\right) W(w),$$

K is a nondegenerate real matrix and ρ_β is a centered Gaussian density operator with the real symmetric covariance matrix β. In [6] it is shown that without loss of generality we can put $K = I$ and consider the POVM

$$M(\mathrm{d}^{2s}z) = D(z)\rho_\beta D(z)^* \frac{\mathrm{d}^{2s}z}{(2\pi)^s} \tag{15}$$

and the corresponding measurement channel $\mathcal{M}$. In our case μ is just the normalized Lebesgue measure on $Z = \mathbb{R}^{2s}$. The fact that the formulas (14) and (15) determine POVM follows from theorem 4.8.3 of [13].

Consider the quadratic Hamiltonian

$$H = R\epsilon R^t, \tag{16}$$

where R is the vector of canonical observables defined by (A3), $\epsilon = \left[\epsilon_{jk}\right]$ is positive definite real symmetric matrix, so that the mean energy of the input state ρ is equal to

$$\mathrm{Tr}\, \rho H = \mathrm{Sp}\, \epsilon\alpha,$$

where Sp denotes trace of $2s \times 2s$-matrices as distinct from the trace of operators in the Hilbert space and

$$\alpha = \mathrm{Re}\, \mathrm{Tr}\, R^t \rho R$$

is the covariance matrix of the state ρ. Then the input energy constraint has the form $\mathrm{Sp}\, \epsilon\alpha \leqslant E$, where E is a positive number. Let us show that in the Gaussian case we are considering,

the conditions of propositions 1 and 2 are fulfilled allowing to compute the *energy-constrained classical capacity* of the channel $\mathcal{M}$.

Lemma 2. *The conditions of proposition 1 are fulfilled for any positive definite quadratic form $H_c(z)$.*

Proof. All such forms are equivalent in the sense $H_{c,1}(z) \asymp H_{c,2}(z)$ for all z, i.e. there exist positive constants k_1, k_2, such that $k_1 H_{c,1}(z) \leqslant H_{c,2}(z) \leqslant k_2 H_{c,1}(z)$. Therefore it is sufficient to prove the inequality

$$\int_{\mathbb{R}^{2s}} H_c(z)\, M(\mathrm{d}^{2s}z) = \int_{\mathbb{R}^{2s}} H_c(z)\, D(z)\rho_\beta D(z)^* \frac{\mathrm{d}^{2s}z}{(2\pi)^s} \leqslant c_1 H + c_2 I \tag{17}$$

where $c_1, c_2 > 0$, and H is given by (16), for at least one such form $H_c(z)$.

Choose a symplectic basis $\{e_j, h_j\}$ associated with the matrix β (see appendix, with α replaced by β) and let the real $2s$-vector z be represented by its coordinates in the decomposition

$$z = \sum_{j=1}^{s} x_j e_j + y_j h_j,$$

so that

$$Rz = \sum_{j=1}^{s} x_j \tilde{q}_j + y_j \tilde{p}_j,$$

where $\tilde{q}_j = Re_j$, $\tilde{p}_j = Rh_j$. Consider the corresponding complex representation of the symplectic space, where z is replaced by the complex s-vector $\mathbf{z}$ with the components $z_j = \frac{1}{\sqrt{2}}\left(x_j + iy_j\right)$ (see e.g. [6]). We will prove the identity

$$\int_{\mathbb{C}^s} \mathbf{z}^* \mathbf{z}\, D(\mathbf{z})\rho_\beta D(\mathbf{z})^* \frac{\mathrm{d}^{2s}\mathbf{z}}{\pi^s} = \tilde{H} + c_2 I. \tag{18}$$

where $\tilde{H} = \frac{1}{2}\sum_{j=1}^{s}\left(\tilde{q}_j^2 + \tilde{p}_j^2\right)$ and $c_2 > 0$. This implies (17) because $\tilde{H}$ is a positive definite quadratic form in R, namely $\tilde{H} = RQR^t$, with $Q = \frac{1}{2}\sum_{j=1}^{s}\left(e_j e_j^t + h_j h_j^t\right) \leqslant c_1 I_{2s}$, the symbol I_{2s} denotes the unit $2s \times 2s$-matrix and $c_1 = \|Q\|$.

The density operator ρ_β admits P-representation

$$\rho_\beta = \int_{\mathbb{C}^s} |\mathbf{w}\rangle \langle \mathbf{w}| \exp\left[-\mathbf{w}^* N_\beta^{-1} \mathbf{w}\right] \frac{\mathrm{d}^{2s}\mathbf{w}}{\pi^s \det N_\beta}$$

where N_β is the diagonal matrix with the entries $N_j > 0, j = 1, \ldots, s$, (see e.g. [13]). The quantity in the left hand side of (18) is the same as

$$\int_{\mathbb{C}^s}\int_{\mathbb{C}^s} \mathbf{z}^* \mathbf{z} \,|\mathbf{z}+\mathbf{w}\rangle \langle \mathbf{z}+\mathbf{w}| \exp\left[-\mathbf{w}^* N_\beta^{-1} \mathbf{w}\right] \frac{\mathrm{d}^{2s}\mathbf{w}}{\pi^s \det N_\beta} \frac{\mathrm{d}^{2s}\mathbf{z}}{\pi^s}.$$

By changing variable $\mathbf{z} + \mathbf{w} \to \mathbf{z}$ and taking into account zero first moments of the Gaussian distribution, we obtain that this is equal to

$$
\int_{\mathbb{C}^s} \int_{\mathbb{C}^s} (\mathbf{z} - \mathbf{w})^* (\mathbf{z} - \mathbf{w}) |\mathbf{z}\rangle \langle \mathbf{z}| \exp\left[-\mathbf{w}^* N_\beta^{-1} \mathbf{w}\right] \frac{d^{2s}\mathbf{w}}{\pi^s \det N_\beta} \frac{d^{2s}\mathbf{z}}{\pi^s}
$$

$$
= \int_{\mathbb{C}^s} \sum_{j=1}^s z_j |\mathbf{z}\rangle \langle \mathbf{z}| \bar{z}_j \frac{d^{2s}\mathbf{z}}{\pi^s} + \int_{\mathbb{C}^s} \mathbf{w}^* \mathbf{w} \exp\left[-\mathbf{w}^* N_\beta^{-1} \mathbf{w}\right] \frac{d^{2s}\mathbf{w}}{\pi^s \det N_\beta} I,
$$

where we used the identity

$$
\int_{\mathbb{C}^s} |\mathbf{z}\rangle \langle \mathbf{z}| \frac{d^{2s}\mathbf{z}}{\pi^s} = I.
$$

Taking into account that $z_j|\mathbf{z}\rangle = \tilde{a}_j|\mathbf{z}\rangle$, where $\tilde{a}_j = \frac{1}{\sqrt{2}}\left(\tilde{q}_j + i\tilde{p}_j\right)$, and using the second moments of the Gaussian distribution, we obtain

$$
\int_{\mathbb{C}^s} \sum_{j=1}^s \tilde{a}_j |\mathbf{z}\rangle \langle \mathbf{z}| \tilde{a}_j^\dagger \frac{d^{2s}\mathbf{z}}{\pi^s} + \left(\operatorname{Sp} N_\beta\right) I = \sum_{j=1}^s \tilde{a}_j \tilde{a}_j^\dagger + \left(\operatorname{Sp} N_\beta\right) I
$$

$$
= \frac{1}{2}\sum_{j=1}^s \left(\tilde{q}_j^2 + \tilde{p}_j^2\right) + \left(s/2 + \operatorname{Sp} N_\beta\right) I.
$$

Putting $c_2 = \left(s/2 + \operatorname{Sp} N_\beta\right)$, we obtain (18). $\qquad\square$

Gaussian measurement channel (15) is covariant with respect to the irreducible (projective) representation $z \to D(z)$ of the additive group $G = Z$ in the sense

$$
D(z)^* M(B) D(z) = M(B - z), \quad z \in Z
$$

for M given by (15), as follows from (A2). Notice that G_0 is trivial in this case. We will now show that the channel satisfies the conditions of the proposition 2, in particular (c) is fulfilled under certain 'threshold condition', in which case the formula (11) holds.

Probability density of outcomes of the observable (15) for the Gaussian input state ρ_α is

$$
p_{\rho_\alpha}(z) = \operatorname{Tr} \rho_\alpha D(z) \rho_\beta D(z)^*
$$

$$
= \int \exp\left(-\frac{1}{2} w^t \alpha w\right) \exp\left(-i w^t z - \frac{1}{2} w^t \beta w\right) \frac{d^{2s}w}{(2\pi)^s}
$$

$$
= \frac{1}{\sqrt{(2\pi)^s \det(\alpha + \beta)}} \exp\left(-\frac{1}{2} z^t (\alpha + \beta)^{-1} z\right).
$$

The Parseval's formula was used here for the Weyl transform (theorem 5.3.3 of [13]).

Denote by $\mathfrak{S}(\alpha)$ the set of all states which have zero first moments, and finite second moments with the covariance matrix α. Take an ensemble π, such that $\bar{\rho}_\pi \in \mathfrak{S}(\alpha)$. Then $p_{\bar{\rho}_\pi}$ is centered probability density with the covariance matrix $\alpha + \beta$, and arguing similarly to the proof of theorem 1 in [6], we get that the maximum of $h\left(p_{\bar{\rho}_\pi}\right)$, equal to

$$
h\left(p_{\rho_\alpha}\right) = \frac{1}{2}\log\det(\alpha + \beta) + C, \tag{19}
$$

where the constant C depends on the normalization of the Lebesgue measure involved in the calculation of the differential entropy (see formula (A8) in appendix), is attained on the Gaussian state $\bar{\rho}_\pi = \rho_\alpha$.

Making additional maximization with respect to Gaussian states with covariance matrix α satisfying the energy constraint $\mathrm{Sp}\,\epsilon\alpha \leqslant E$, we obtain

$$\max_{\rho:\mathrm{Tr}\,\rho H \leqslant E} h\left(p_\rho\right) = \max_{\alpha:\mathrm{Sp}\,\alpha\epsilon \leqslant E} \frac{1}{2} \log \det\left(\alpha + \beta\right) + C,$$

and the maximizer is a centered Gaussian state ρ_E^0 with the covariance matrix

$$\alpha_E^0 = \arg \max_{\alpha:\mathrm{Sp}\,\alpha\epsilon \leqslant E} \det\left(\alpha + \beta\right). \tag{20}$$

Thus, (a) of proposition 2 follows.

The statement of (b) follows from the results concerning the minimal output entropy. The result of the paper [15] (proposition 4; see also [6]) concerning the minimal output entropy of the Gaussian measurement channel implies that the minimizer ρ_0 can be taken as the vacuum state related to the complex structure J_β (see appendix). Substituting $\alpha = \frac{1}{2}\Delta J_\beta$ into (19), we get

$$\min_\rho h\left(p_\rho\right) = h\left(p_{\rho_0}\right) = \frac{1}{2} \log \det\left(\beta + \frac{1}{2}\Delta J_\beta\right) + C.$$

The condition n. 3 is fulfilled provided

$$\alpha_E^0 \geqslant \frac{1}{2}\Delta J_\beta. \tag{21}$$

Indeed, in this case

$$\rho_E^0 = \int_{\mathbb{R}^{2s}} D(z)\rho_0 D(z)^* \pi_E^0(\mathrm{d}z),$$

where $\pi_E^0(\mathrm{d}z)$ is the centered Gaussian distribution on Z with the covariance matrix $\alpha_E^0 - \frac{1}{2}\Delta J_\beta$. One can check this by comparing the quantum characteristic functions of both sides. The matrix inequality (21) is an analogue of the 'threshold condition' for the multi-mode quantum Gaussian channel [12] (in the case of one mode inequalities of this kind appeared in [16–18]).

Thus we have proved the following result extending theorem 1 of [6] to the case without global gauge symmetry.

Theorem 1. *Let $\widetilde{\mathcal{M}}$ be the measurement channel corresponding to the Gaussian POVM (14). Assume that α_E^0, given by the formula (20), and β satisfy the condition (21). Then*

$$C(\widetilde{\mathcal{M}}, H, E) = \frac{1}{2} \log \det\left(\alpha_E^0 + \beta\right) - \frac{1}{2} \log \det\left(\beta + \frac{1}{2}\Delta J_\beta\right)$$

$$= \frac{1}{2} \log \det\left[I + \left(\alpha_E^0 - \frac{1}{2}\Delta J_\beta\right)\left(\beta + \frac{1}{2}\Delta J_\beta\right)^{-1}\right], \tag{22}$$

which is attained on the ensemble of J_β-coherent states $D(z)\rho_0 D(z)^*$ (see appendix), where z has the centered Gaussian probability distribution with the covariance matrix $\alpha_E^0 - \frac{1}{2}\Delta J_\beta$.

Finding the covariance matrix (20) is a separate finite-dimensional optimization problem which can be solved analytically in some special cases.

Example. Take the energy operator of the signal mode $H = (q^2 + p^2)/2$ with the corresponding complex structure $J_H = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. We consider the POVM described by the formula (15) where the covariance matrix of the quantum Gaussian noise is

$$\beta = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}; \quad \beta_1 \beta_2 \geqslant \frac{1}{4}.$$

In quantum optics this would correspond to the heterodyne measurement of the signal mode with the squeezed quantum noise from the local oscillator.

The complex structure of the measurement noise is

$$J_\beta = \begin{bmatrix} 0 & -\sqrt{\beta_2/\beta_1} \\ \sqrt{\beta_1/\beta_2} & 0 \end{bmatrix},$$

which does not commute with J_H unless $\beta_1 = \beta_2$. The covariance matrix of the squeezed vacuum is (see appendix)

$$\frac{1}{2}\Delta J_\beta = \frac{1}{2}\begin{bmatrix} \sqrt{\beta_1/\beta_2} & 0 \\ 0 & \sqrt{\beta_2/\beta_1} \end{bmatrix}.$$

and

$$\beta + \frac{1}{2}\Delta J_\beta = \begin{bmatrix} \beta_1 + \frac{1}{2}\sqrt{\beta_1/\beta_2} & 0 \\ 0 & \beta_2 + \frac{1}{2}\sqrt{\beta_2/\beta_1} \end{bmatrix},$$

so that $\det\left(\beta + \frac{1}{2}\Delta J_\beta\right) = \left(\sqrt{\beta_1 \beta_2} + 1/2\right)^2$, hence the second term in (22) is $-\log\left(\sqrt{\beta_1 \beta_2} + 1/2\right)$.

To compute the first term, we can restrict to diagonal input covariance matrices

$$\alpha = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}, \quad \alpha_1 + \alpha_2 \leqslant 2E, \alpha_1 \alpha_2 \geqslant \frac{1}{4}.$$

The matrix

$$\alpha + \beta = \begin{bmatrix} \beta_1 + \alpha_1 & 0 \\ 0 & \beta_2 + \alpha_2 \end{bmatrix}$$

has the determinant $(\beta_1 + \alpha_1)(\beta_2 + \alpha_2)$, so that the maximized expression is $\log\sqrt{(\beta_1 + \alpha_1)(\beta_2 + \alpha_2)}$. Since $\log x$ is increasing function, we have to maximize $(\beta_1 + \alpha_1)(\beta_2 + \alpha_2)$ under the constraints $\alpha_1 + \alpha_2 \leqslant 2E, \quad \alpha_1 \alpha_2 \geqslant \frac{1}{4}$. The first constraint gives the values

$$\alpha_1^0 = E + (\beta_2 - \beta_1)/2, \quad \alpha_2^0 = E - (\beta_2 - \beta_1)/2,$$

corresponding to the maximal value of the first term in (22)

$$\log\frac{1}{2}\left(2E + (\beta_1 + \beta_2)\right).$$

The second constraint will be automatically fulfilled provided we impose the condition (21) which amounts to

$$\alpha_1^0 \geqslant \frac{1}{2}\sqrt{\beta_1/\beta_2}, \quad \alpha_2^0 \geqslant \frac{1}{2}\sqrt{\beta_2/\beta_1},$$

or

$$E \geqslant \frac{1}{2}\left(\max\left\{\sqrt{\beta_1/\beta_2}, \sqrt{\beta_2/\beta_1}\right\} + |\beta_2 - \beta_1|\right).$$

Under this condition

$$C(\mathcal{M}; H, E) = \log\left(E + (\beta_1 + \beta_2)/2\right) - \log\left(\sqrt{\beta_1\beta_2} + 1/2\right)$$

$$= \log\left(\frac{2E + (\beta_1 + \beta_2)}{2\sqrt{\beta_1\beta_2} + 1}\right).$$

Acknowledgments

Appendix

In this section we systematically use notations and some results from the book [14] where further references are given. Consider a finite-dimensional symplectic space (Z, Δ) with $Z = \mathbb{R}^{2s}$,

$$\Delta = \text{diag}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}_{j=1,\dots,s}. \tag{A1}$$

Let $\mathcal{H}$ be the space of an irreducible representation $z \to W(z); z \in Z$ of the canonical commutation relations

$$W(z)W(z') = \exp\left[-\frac{i}{2}z^t\Delta z'\right]W(z + z'). \tag{A2}$$

Here $W(z) = \exp iRz$ are the unitary Weyl operators, where

$$Rz = \sum_{j=1}^{s}(x_jq_j + y_jp_j), \tag{A3}$$

$z = [x_j, y_j]_{j=1,\dots,s}^t$, are the canonical observables of the quantum system.

Operator J in (Z, Δ) is called *operator of complex structure* if

$$J^2 = -I_{2s}, \tag{A4}$$

where I_{2s} is the identity operator in Z, and it is Δ-positive in the sense that

$$\Delta J = -J^t\Delta, \quad \Delta J \geqslant 0. \tag{A5}$$

A centered Gaussian state ρ_α on $\mathfrak{B}(\mathcal{H})$ is determined by its covariance matrix $\alpha =$ Re Tr $R^t \rho R$ which is a real symmetric $2s \times 2s$-matrix satisfying

$$\alpha \geqslant \pm \frac{i}{2}\Delta. \tag{A6}$$

This state is pure if and only if $\alpha = \frac{1}{2}\Delta J$. It is called J-vacuum and denoted ρ_0. The non-centered pure states $D(z)\rho_0 D(z)^*$ are called J-coherent states (see section 12.3.2 of [14]).

Consider the operator $A = \Delta^{-1}\alpha$. The operator A is skew-symmetric in the Euclidean space (Z, α) with the scalar product $\alpha(z, z') = z^t \alpha z'$. According to a theorem from linear algebra, there is an orthogonal basis $\{e_j, h_j\}$ in (Z, α) and positive numbers $\{\alpha_j\}$ such that

$$Ae_j = \alpha_j h_j; \quad Ah_j = -\alpha_j e_j.$$

Equation (A6) implies $N_j \equiv \alpha_j - 1/2 \geqslant 0$. Choosing the normalization $\alpha(e_j, e_j) = \alpha(h_j, h_j) = \alpha_j$ gives the symplectic basis in (Z, Δ) with the required properties.

There is at least one operator of complex structure, commuting with the operator $A = \Delta^{-1}\alpha$, namely, the orthogonal operator J_α from the polar decomposition

$$A = |A| J_\alpha = J_\alpha |A| \tag{A7}$$

in the Euclidean space (Z, α). The action of J_α in the symplectic basis is given by the formula

$$J_\alpha e_j = h_j, \quad J_\alpha h_j = -e_j.$$

In [14] several formulas for the entropy of state ρ_α were given. While we do not need these in the present paper, let us show just another convenient formula involving the complex structure J_α:

$$H(\rho_\alpha) = \frac{1}{2}\mathrm{Sp}\, g\left(\alpha - 1/2\Delta J_\alpha\right),$$

where $g(x) = (x + 1)\log(x + 1) - x \log x$. Indeed, starting from the relation (12.108) in [14], we have

$$H(\rho_\alpha) = \sum_{j=1}^{s} g\left(N_j\right) = \sum_{j=1}^{s} g\left(\alpha_j - 1/2\right)$$

$$= \frac{1}{2}\mathrm{Sp}\, g\left(\tilde{\alpha} - 1/2\Delta\tilde{J}_\alpha\right) = \frac{1}{2}\mathrm{Sp}\, g\left(\alpha - 1/2\Delta J_\alpha\right),$$

where the last step follows from the fact that

$$\alpha - 1/2\Delta J_\alpha = T^t\left(\tilde{\alpha} - 1/2\Delta\tilde{J}_\alpha\right) T,$$

T is the transition matrix between the symplectic bases, $\det T = 1$.

We will need the formula for the differential entropy of a nondegenerate multidimensional Gaussian probability distribution p_α with the covariance matrix α:

$$h(p_\alpha) = \frac{1}{2}\log \det \alpha + C = \frac{1}{2}\mathrm{Sp}\, \log \alpha + C, \tag{A8}$$

where the constant C depends on the normalization of the Lebesgue measure involved in the definition of the differential entropy (see [9]).

ORCID iDs

A S Holevo ⬤ https://orcid.org/0000-0001-5699-521X

References

[1] Dall'Arno M, D'Ariano G M and Sacchi M F 2011 Informational power of quantum measurements *Phys. Rev.* A **83** 062304

[2] Hall M J W 1997 Quantum information and correlation bounds *Phys. Rev.* A **55** 1050–2947

[3] Holevo A S 2012 Information capacity of quantum observable *Probl. Inf. Transm.* **48** 1–10

[4] Oreshkov O, Calsamiglia J, Munoz-Tapia R and Bagan E 2011 Optimal signal states for quantum detectors *New J. Phys.* **13** 073032

[5] Kuznetsova A A and Holevo A S 2015 Coding theorems for hybrid channels. II *Theor. Probab. Appl.* **59** 145–54

[6] Holevo A S 2019 Gaussian maximizers for quantum Gaussian observables and ensembles (arXiv:1908.03038 [math-ph])

[7] Holevo A S 2008 Entanglement-breaking channels in infinite dimensions *Probl. Inf. Transm.* **44** 171–84

[8] Holevo A S and Shirokov M E 2005 Continuous ensembles and the capacity of infinite-dimensional quantum channels *Theor. Probab. Appl.* **50** 86–98

[9] Cover T M and Thomas J A 1996 *Elements of Information Theory* 2nd edn (New York: Wiley)

[10] Holevo A S 2004 Entanglement-assisted capacities of constrained quantum channels *Theor. Probab. Appl.* **48** 243–55

[11] Dobrushin R L 1959 General formulation of Shannon theorem in information theory *Russ. Math. Surv.* **14** 3–104

[12] Holevo A S 2016 On the constrained classical capacity of infinite-dimensional covariant channels *J. Math. Phys.* **57** 15203

[13] Holevo A S 2011 *Probabilistic and Statistical Aspects of Quantum Theory* 2nd English edn (Pisa: Edizioni Della Normale)

[14] Holevo A S 2019 *Quantum Systems Channels Information: A Mathematical Introduction* 2nd edn (Berlin: De Gruyter)

[15] Giovannetti V, Holevo A S and Mari A 2015 Majorization and additivity for multimode bosonic Gaussian channels *Theor. Math. Phys.* **182** 284–93

[16] Lupo C, Pilyavets O V and Mancini S 2009 Capacities of lossy bosonic channel with correlated noise *New J. Phys.* **11** 063023

[17] Lupo C, Pirandola S, Aniello P and Mancini S 2011 On the classical capacity of quantum Gaussian channels *Phys. Scr.* **T143** 014016

[18] Schäfer J, Karpov E, García-Patrón R, Pilyavets O V and Cerf N J 2013 Equivalence relations for the classical capacity of single-mode Gaussian quantum channels *Phys. Rev. Lett.* **111** 030503