CODING THEOREMS FOR HYBRID CHANNELS. II

Kuznetsova A. A.† Holevo A. S.‡

Abstract

The present work continues investigation of the capacities of measurement (quantum-classical) channels in the most general setting, initiated in [10]. The proof of coding theorems is given for the classical capacity and entanglement-assisted classical capacity of the measurement channel with arbitrary output alphabet, without assuming that the channel is given by a bounded operator-valued density.

1 Introduction

The present work continues investigation of the capacities of measurement channels in the most general setting, initiated in [10]. The proof of coding theorems is given for the classical capacity (theorem 1) and entanglement-assisted classical capacity (theorem 2) of the measurement channel with arbitrary output alphabet under the minimal regularity assumptions. The statement of theorem 2 was proved previously in [10] under additional assumption that the channel is given by a bounded operator-valued density. In the present work we relax this restriction by using a generalization of the Radon-Nikodym theorem for probability operator-valued measures [6]. The result obtained is illustrated by an example of homodyne measurement in quantum optics.

†Work partially supported by RFBR (grant No 12-01-00319) and Russian Quantum Center.

†N. E. Baumann MHTU, Moscow, Russia (kuznetsova.a.a@bk.ru).

‡Steklov Mathematical Institute, RAS, Moscow, Russia (holevo@mi.ras.ru).
We remark that the entanglement-assisted classical capacity was studied by a number of authors under the names purification capacity, measurement strength, forward classical communication cost. In the recent paper [2], where one can find further references, its alternative interpretation is developed. It is shown that a (finite-dimensional) measurement channel can be asymptotically simulated by transmission of a classical message of the size equal to the maximal entropy reduction, assisted with sufficient classical correlation between the input and the output. The result can be considered as a quantum reverse Shannon theorem in which entanglement and quantum channel are replaced, correspondingly, by classical correlation and classical channel.

2 Preliminaries

Let \( \mathcal{H} \) be a separable Hilbert space. We use the following notations: \( \mathcal{B}(\mathcal{H}) \) is the algebra of all bounded operators, \( \mathcal{S}(\mathcal{H}) \) is the space of trace-class operators in \( \mathcal{H} \), \( \mathcal{S}(\mathcal{H}) \) is its convex subset of density operators (i.e. positive operators with unit trace), called also quantum states.

We introduce the measure space \((\Omega, \mathcal{F}, \mu)\), where \(\Omega\) is a complete separable metric space, \(\mathcal{F}\) is a \(\sigma\)-algebra of its subsets, \(\mu\) is a \(\sigma\)-finite measure on \(\mathcal{F}\). A hybrid (classical-quantum) system is described by von Neumann algebra \( \mathcal{L} = L^\infty(\Omega, \mathcal{F}, \mu; \mathcal{B}(\mathcal{H})) \), consisting of weakly measurable, essentially bounded functions \( X(\omega), \omega \in \Omega \) with values in \( \mathcal{B}(\mathcal{H}) \). Consider the predual joint space \( \mathcal{L}^* = L_1(\Omega, \mathcal{F}, \mu; \mathcal{S}(\mathcal{H})) \), the elements of which are measurable functions \( S = \{S(\omega)\} \) with values in \( \mathcal{S}(\mathcal{H}) \), integrable with respect to the measure \( \mu \). An element \( S = \{S(\omega)\} \in \mathcal{L}^* \) such that

\[
S(\omega) \geq 0 \pmod{\mu}, \quad \int_\Omega \text{Tr} S(\omega) \mu(d\omega) = 1,
\]

is called state on the algebra \( \mathcal{L} \). In notations of entropic characteristics of hybrid systems we will use the index “cq”, of classical and quantum systems — the indices “c” and “q” correspondingly.

Following [4], we introduce the notions of entropy and relative entropy of cq-states. Concerning the definitions and properties of quantum entropies see e. g. [8].

**Definition 1.** The entropy of a cq-state \( S \) is defined by the relation

\[
H_{cq}(S) = \int_\Omega H_q(S(\omega)) \mu(d\omega),
\]
where \( H_q(S) = -\text{Tr} S \log S \) is the von Neumann entropy of positive operator \( S \in \mathcal{S}(\mathcal{H}) \).

Note that

\[
H_{cq}(S) = H_c(p) + \int_\Omega p(\omega) H_q(\hat{S}(\omega)) \mu(d\omega), \tag{1}
\]

where \( p(\omega) = \text{Tr} S(\omega), \hat{S}(\omega) = (p(\omega))^{-1} S(\omega), H_c(p) \) is the differential entropy of the probability distribution with the density \( p(\omega) \) with respect to the measure \( \mu \).

**Definition 2.** The relative entropy of cq-states \( S_1, S_2 \) is defined by the relation

\[
H_{cq}(S_1 \parallel S_2) = \int_\Omega H_q(S_1(\omega) \parallel S_2(\omega)) \mu(d\omega),
\]

where

\[
H_q(S_1(\omega) \parallel S_2(\omega)) = \text{Tr} S_1(\omega)(\log S_1(\omega) - \log S_2(\omega))
\]

is the quantum relative entropy.

To describe measurement channels we will need the following definition.

**Definition 3.** Probability operator-valued measure (POVM) on \( \Omega \) is a family \( M = \{M(A), A \in \mathcal{F}\} \) of bounded Hermitian operators in \( \mathcal{H} \), satisfying the conditions:

1) \( M(A) \geq 0, A \in \mathcal{F} \);

2) \( M(\Omega) = I \), where \( I \) is the unit operator in \( \mathcal{H} \);

3) for arbitrary countable decomposition \( A = \bigcup_i A_i (A_i \cap A_j = \emptyset, i \neq j) \), the relation \( M(A) = \sum_i M(A_i) \) holds in the sense of weak convergence of operators.

POVM defines a quantum observable with values in \( \Omega \). The probability distribution of observable \( M \) in the state \( S \) is given by the formula

\[
P_S(A) = \text{Tr} SM(A), \quad A \in \mathcal{F}. \tag{2}
\]

For brevity, we sometimes write \( P_S(d\omega) = \text{Tr} SM(d\omega) \).

If POVM \( M(d\omega) \) is defined by the density \( P(\omega) \) with respect to scalar \( \sigma \)-finite measure \( \mu \), where \( P(\omega) \) is a uniformly bounded (with respect to the operator norm) weakly measurable operator-valued function, then its probability distribution has the density \( p_S(\omega) = \text{Tr} SP(\omega) \) with respect to the measure \( \mu \). This case is studied in [10].
In the general case the following lemma holds (a generalization of the Radon-Nikodym theorem for POVM [6]).

**Lemma 1.** For an arbitrary POVM on a separable metric space \( \Omega \) there exist a dense subspace \( \mathcal{D} \in \mathcal{H} \), a \( \sigma \)-finite measure \( \mu \) on \( \Omega \), a countable set of Borel functions \( \omega \rightarrow a_k(\omega) \) where for almost all \( \omega \) the \( a_k(\omega) \) are linear functionals on \( \mathcal{D} \), satisfying the conditions

\[
\int_\Omega \sum_k |\langle a_k(\omega)|\psi\rangle|^2 \mu(d\omega) = \|\psi\|^2, \quad \psi \in \mathcal{D},
\]  

\[ \langle \psi|M(A)|\psi\rangle = \int_A \sum_k |\langle a_k(\omega)|\psi\rangle|^2 \mu(d\omega), \quad \psi \in \mathcal{D}. \]  

In [6] it is shown that for \( \mathcal{D} \) one can take \( \text{lin} \{\varphi_i\} \) — the linear span of a fixed orthonormal basis \( \{\varphi_i\} \).

**Lemma 2.** For arbitrary observable \( M(d\omega) \) with values in \( \Omega \) and a density operator \( S \in \mathcal{S}(\mathcal{H}) \), the probability distribution \( P_S(d\omega) = \text{Tr} SM(d\omega) \) has density \( p_S(\omega) \) with respect to measure \( \mu \).

**Proof.** Consider the spectral decomposition of the state \( S \):

\[
S = \sum_{i=1}^{\infty} \lambda_i |\varphi_i\rangle \langle \varphi_i|. \]  

Apply lemma 1 with \( \mathcal{D} = \text{lin}\{\varphi_i\} \). For all \( A \in \mathcal{F} \) the equality holds

\[
P_S(A) \equiv \text{Tr} SM(A) = \int_A p_S(\omega) \mu(d\omega),
\]  

where

\[
p_S(\omega) = \sum_{i=1}^{\infty} \lambda_i \sum_k |\langle a_k(\omega)|\varphi_i\rangle|^2
\]  

is a nonnegative integrable function by the condition (3) and the spectral decomposition (5). The lemma is proved.

Let us fix an orthonormal system \( \{e_k\} \) in \( \mathcal{H} \). According to the same conditions (3) and (5), the relation

\[
\tilde{S}(\omega) = (p_S(\omega))^{-1} \sum_{i=1}^{\infty} \lambda_i \sum_{j,k} |e_k\rangle \langle a_k(\omega)|\varphi_i\rangle \langle a_j(\omega)|\varphi_i\rangle \langle e_j|
\]  

is a nonnegative function on \( \Omega \).
for $P_S$-almost all $\omega$ defines a density operator in $\mathcal{H}$, which we will call posterior state. The meaning of this term is that under certain conditions the operator $\hat{S}(\omega)$ describes state of the quantum system after measurement of observable $M$, which resulted with the outcome $\omega$ [8].

Following [7], define the entropy reduction by the relation

$$ER(S, M) = H_q(S) - \int_{\Omega} p(\omega) H_q(\hat{S}(\omega)) \mu(d\omega),$$

which is consistent provided $H_q(S) < \infty$. We mention the following approximation properties. Consider a sequence of states $S_n = \sum_{i=1}^{n} \tilde{\lambda}_i |\varphi_i\rangle\langle \varphi_i|$, where $\tilde{\lambda}_i = (\sum_{k=1}^{n} \lambda_k)^{-1}\lambda_i$. By lemma 4 of the paper [11], the above sequence $S_n$ satisfies the condition

$$\lim_{n \to \infty} H_q(S_n) = H_q(S) < \infty. \quad (9)$$

According to the theorem 2 from [7], this implies

$$\lim_{n \to \infty} ER(S_n, M) = ER(S, M) < \infty. \quad (10)$$

3 The classical capacity of a measurement channel

**Definition 4.** Let $M$ be a POVM, $P_S$ — its probability distribution in the state $S$, which is given by the formula (2). Measurement channel $\mathcal{M}$ is an affine map $S \to P_S(d\omega)$ of the convex set of quantum states $\mathcal{S}(\mathcal{H})$ into the set of probability distributions on $\Omega$.

To apply the method of block coding, we need to define the $n$-th degree $\mathcal{M}^\otimes n$ of the channel $\mathcal{M}$. Let $\mathcal{H}^\otimes n$ be the $n$-th tensor degree of the Hilbert space $\mathcal{H}$ and let $(\Omega^\times n, \mathcal{F}^\times n)$ be the product of $n$ copies of the measurable space $(\Omega, \mathcal{F})$. The channel $\mathcal{M}^\otimes n$ is defined by the observable $M^\otimes n$ with values in $\Omega^\times n$ such that

$$M^\otimes n(A_1 \times \cdots \times A_n) = M(A_1) \otimes \cdots \otimes M(A_n).$$

By using an analog of the extension theorem for POVM, one can show that this relation defines uniquely all the values $M^\otimes n(A^{(n)})$, $A^{(n)} \in \mathcal{F}^\times n$. 

In the case of infinite-dimensional $\mathcal{H}$ one usually introduces a constraint onto the input states of the channel (otherwise the capacities are infinite as a rule). Let $F$ be a positive selfadjoint (in general unbounded) operator in the space $\mathcal{H}$, with the spectral decomposition $F = \int_0^\infty x \, dE(x)$, where $E(x)$ is the spectral function. We introduce the subset of states
\[ \mathcal{A}_E = \{ S \in \mathfrak{S}(\mathcal{H}) : \text{Tr} SF \leq E \}, \] (11)
where $E$ is a positive constant, and the trace in (11) is understood as the integral $\int_0^\infty x \, d(\text{Tr} SE(x))$ (for more detail see [3]). Notice that if the operator $F$ satisfies the condition
\[ \text{Tr} \exp(-\beta F) < \infty \quad \beta > 0, \] (12)
then $H_q(S) < \infty$ for all $S$ such that $\text{Tr} SF \leq E$ (see [3]). The corresponding constraint for the channel $\mathcal{M}^\otimes n$ is determined by the operator
\[ F^{(n)} = F \otimes I \otimes \cdots \otimes I + \cdots + I \otimes I \otimes \cdots \otimes F. \]

Denote
\[ \mathcal{A}_E^{(n)} = \{ S^{(n)}_i \in \mathfrak{S}(\mathcal{H}^\otimes n) : \text{Tr} S^{(n)}_i F^{(n)} \leq nE \}. \] (13)

**Definition 5.** The code of length $n$ and size $N$ is a pair $(\Sigma^{(n)}, V^{(n)})$, where:
1) $\Sigma^{(n)} = \{ S^{(n)}_i, \ i = 1, \ldots, N \}$ is a family of states from $\mathcal{A}_E^{(n)}$;
2) $V^{(n)} = \{ V_j, \ j = 0, 1, \ldots, N \}$ is a decomposition of the space $\Omega^{\times n}$.

**Definition 6.** The average error probability of the code $(\Sigma^{(n)}, V^{(n)})$ is the quantity
\[ \bar{u}(\Sigma^{(n)}, V^{(n)}) = \frac{1}{N} \sum_{j=1}^{N} \left( 1 - \text{Tr} S^{(n)}_j M^{\otimes n}(V_j) \right). \] (14)

We denote by $u(n, N)$ the greatest lower bound of the quantity $\bar{u}(\Sigma^{(n)}, V^{(n)})$ with respect to all codes of length $n$ and size $N$.

**Definition 7.** We call the classical capacity $C(\mathcal{M}, \mathcal{A}_E)$ of the measurement channel $\mathcal{M}$ with the constraint (13) the supremum of all achievable rates i.e. the values $R > 0$, satisfying the condition
\[ \lim_{n \to \infty} u(n, 2^{nR}) = 0. \] (15)

We call by ensemble of states a finite probability distribution $\pi = \{ \pi_x ; S_x \}$ on the set of states $\mathfrak{S}(\mathcal{H})$, ascribing probabilities $\pi_x$ to certain states $S_x$. The
average state of ensemble is defined as: \( \overline{S}_\pi = \sum_x \pi_x S_x \). Let us denote \( \mathcal{P}_E \) the set of ensembles \( \pi \) such that \( \overline{S}_\pi \in \mathcal{A}_E \); similarly, we denote \( \mathcal{P}_E^{(n)} \) the set of ensembles \( \pi^{(n)} \) in \( \mathcal{G}(\mathcal{H}^\otimes n) \), the average state of which satisfying the condition (13).

For given measurement channel \( \mathcal{M} \) and ensemble \( \pi \) define the quantity
\[
I(\pi, \mathcal{M}) = \sum_x \pi_x \int_\Omega p(\omega|x) \log \frac{p(\omega|x)}{p(\omega)} \mu(d\omega).
\] (16)

Here \( p(\omega) \) and \( p(\omega|x) \) are the probability densities of the distributions \( \overline{P}(d\omega) = \text{Tr} \overline{S}_\pi \mathcal{M}(d\omega) \) and \( P_x(d\omega) = \text{Tr} S_x \mathcal{M}(d\omega) \) correspondingly. The quantity \( I(\pi, \mathcal{M}) \) is the Shannon mutual information between the discrete random variable \( X \), having the probability distribution \( \{\pi_x\} \) and the random variable \( \omega \) with conditional probability density \( p(\omega|x) \), defined via lemma [2] Notice that there is a representation of the quantity \( I(\pi, \mathcal{M}) \) as a supremum over decompositions \( \mathcal{V} = \{V_i\} \) of the output space \( \Omega \) (cf. [3] formula (1.2.3)):
\[
I(\pi, \mathcal{M}) = \sup_{\mathcal{V}} \left( \sum_i \sum_x \pi_x P_x(V_i) \log \frac{P_x(V_i)}{P(V_i)} \right). \tag{17}
\]
The quantity under supremum is equal to \( I(\pi, \mathcal{M}_V) \), where \( \mathcal{M}_V \) is the measurement channel, corresponding to the discrete observable \( \{M(V_i)\} \).

**Theorem 1.** The classical capacity of the measurement channel \( \mathcal{M} \) with the constraint (13) is given by the relation
\[
C(\mathcal{M}, \mathcal{A}_E) = \sup_{\pi \in \mathcal{P}_E} I(\pi, \mathcal{M}). \tag{18}
\]

**Proof.** Denote
\[
C_n = \sup_{\pi^{(n)} \in \mathcal{P}_E^{(n)}} I(\pi^{(n)}, \mathcal{M}_E^{\otimes n}).
\]
We need to show that
\[
C(\mathcal{M}, \mathcal{A}_E) = C_1.
\]
Let us first establish the additivity property \( C_n = nC_1 \).

For a fixed decomposition \( \mathcal{V} = \{V_i\} \), the measurement channel \( \mathcal{M}_V \) is embedded into quantum entanglement-breaking channel (see e.g. [4]), therefore according to [3] its capacity is given by the expression
\[
C(\mathcal{M}_V, \mathcal{A}_E) = \sup_{\pi \in \mathcal{P}_E} I(\pi, \mathcal{M}_V), \tag{19}
\]
and has the additivity property (see [6])

\[ C(\mathcal{M}^\otimes_n, \mathcal{A}_E^n) = nC(\mathcal{M}, \mathcal{A}_E). \]

Notice that similarly to (19), the left-hand side is equal to \( \sup_{\pi(n) \in \mathcal{P}_E} I(\pi(n), \mathcal{M}_V^\otimes n) \), so that

\[ \sup_{\pi(n) \in \mathcal{P}_E} I(\pi(n), \mathcal{M}_V^\otimes n) = n \sup_{\pi \in \mathcal{P}_E} I(\pi, \mathcal{M}_V). \]  

(20)

By using a result of R. L. Dobrushin (theorem 2.2 in [5]), we have

\[ I(\pi(n), \mathcal{M}_V^\otimes n) = \sup_{\mathcal{V}} I(\pi(n), \mathcal{M}_V^\otimes n), \]

because the supremum in the right-hand side is equal to the supremum of the information with respect to decompositions of the space \( \Omega^\times n \) of special form, consisting of products \( V_1 \times \cdots \times V_n \) of the sets from the decomposition \( \mathcal{V} \). The class of all such products has the ordering property that is required for validity of theorem 2.2 in [5]. Hence

\[ C_n = \sup_{\mathcal{V}} \sup_{\pi(n) \in \mathcal{P}_E} I(\pi(n), \mathcal{M}_V^\otimes n) = \sup_{\mathcal{V}} \sup_{\pi(n) \in \mathcal{P}_E} I(\pi(n), \mathcal{M}_V^\otimes n) \]

\[ = n \sup_{\mathcal{V} \pi \in \mathcal{P}_E} I(\pi, \mathcal{M}_V) = n \sup_{\mathcal{V} \pi \in \mathcal{P}_E} I(\pi, \mathcal{M}_V) \]

\[ = n \sup_{\pi \in \mathcal{P}_E} I(\pi, \mathcal{M}) = nC_1, \]

where we used (20) in the third equality.

Now let us prove the inequality \( C(\mathcal{M}, \mathcal{A}_E) \leq C_1 \). Without loss of generality we can suppose that \( C_1 < \infty \). Let \( R > C_1 \). By applying Fano’s inequality, we obtain similarly to the relation (10.19) in [8]

\[ u(n, 2^{nR}) \geq 1 - \frac{C_n}{nR} - \frac{1}{nR} = 1 - \frac{C_1}{R} - \frac{1}{nR}, \]

where in the second equality we used the additivity \( C_n = nC_1 \). Therefore \( \liminf_{n \to \infty} u(n, 2^{nR}) > 0 \), and hence \( C(\mathcal{M}, \mathcal{A}_E) \leq C_1 \).

For the proof of the converse inequality we note that

\[ C(\mathcal{M}, \mathcal{A}_E) \geq C(\mathcal{M}_V, \mathcal{A}_E) = \sup_{\pi \in \mathcal{P}_E} I(\pi, \mathcal{M}_V) \]

for arbitrary decomposition \( \mathcal{V} \). By taking supremum over the decompositions \( \mathcal{V} \), we obtain \( C(\mathcal{M}, \mathcal{A}_E) \geq \sup_{\pi \in \mathcal{P}_E} I(\pi, \mathcal{M}) = C_1 \). The theorem 1 is proved.
4 Entanglement-assisted capacity of a measurement channel

Consider the following protocol of classical information transmission through the measurement channel $M$. Transmitter $A$ and receiver $B$ are in the pure entangled state $S_{AB} = |\psi\rangle\langle\psi|$, where $|\psi\rangle = \sum_j c_j |e_j\rangle \otimes |\tilde{e}_j\rangle$, satisfying the condition $H_q(S_A) = H_q(S_B) < \infty$.

Let $X$ be a finite alphabet, and the classical signal $x \in X$ appears with probability $\pi_x$. The party $A$ performs encoding $x \rightarrow E_x$, and then sends its part of the resulting common state via the channel $M$. Thus party $B$ has at its disposal the hybrid system $\Omega_B$, where $\Omega$ is the classical system at the output of the measurement channel. After the measurement of observable $M(d\omega)$, the state in the hybrid system is described in the following way:

$$\sigma_x(d\omega) = \sum_{j,k} c_j c_k \text{Tr} [E_x(|e_j\rangle \langle e_k|)M(d\omega)] |e_j\rangle \langle e_k|.$$ 

Then the party $B$ may perform measurement of an observable in the system $\Omega_B$, extracting in this way information about the signal $x$.

With the block coding, the encoded states transmitted through the channel $M^\otimes n \otimes I_B^\otimes n$, have the form

$$S^{(n)}_{\alpha} = (E^{(n)}_\alpha \otimes I_B^\otimes n)[S^{(n)}_{AB}], \quad (21)$$

where $S^{(n)}_{AB}$ is pure entangled state for $n$ copies of the system $AB$, satisfying the condition $H(S^{(n)}_B) < \infty$, $\alpha$ is the classical message (e.g. a word in an alphabet $X$), $\alpha \rightarrow E^{(n)}_\alpha$ are the encodings for $n$ copies of the system $A$. The input states of the channel $M^\otimes n$ are subject to the constraint (13), which is equivalent to similar constraint for the channel $M^\otimes n \otimes I_B^\otimes n$ with the operators $F^{(n)} \otimes I_B^\otimes n$.

For the channel $M$ with the input constraint (13) we consider the quantity

$$C^{(n)}_{ca}(M^\otimes n, A^{(n)}_E) = \sup_{(\pi^{(n)}_{\alpha}, S^{(n)}_{\alpha})} \chi_{cq} \left( \{\pi^{(n)}_{\alpha}\}; \{M^\otimes n \otimes I_B^\otimes n)S^{(n)}_{\alpha}\} \right), \quad (22)$$

where

$$\chi_{cq} (\{\pi_x\}; \{S_x\}) = H_{cq} \left( \sum_x \pi_x S_x \right) - \sum_x \pi_x H_{cq} (S_x), \quad S_x \in \mathcal{L}. $$
and the supremum is taken over all state ensembles of the form (21), satisfying the condition
\[ \sum_\alpha \pi_\alpha^{(n)} \text{Tr} S_\alpha^{(n)} (F^{(n)} \otimes I_B^{\otimes n}) \leq nE. \]

The classical entanglement-assisted capacity for the quantum-classical channel \( \mathcal{M} \) with the constraint (11) is defined by the relation
\[ C_{ea}(\mathcal{M}, A_E) = \lim_{n \to \infty} \frac{1}{n} C_{ea}^{(n)} (\mathcal{M}^{\otimes n}, A_E^{(n)}). \]

**Theorem 2.** Let \( \mathcal{M} \) be an arbitrary measurement channel with the input constraint (13). Assume, that the operator \( F \) satisfies the condition (12), and the channel \( \mathcal{M} \) satisfies the condition
\[ \sup_{S, A: \text{Tr} S A F \leq E} \text{Tr} S A F \leq E H_c (p S A) < \infty, \] (23)
where \( H_c (p S A) \) is the classical differential entropy of the probability density of the output distribution of the channel \( \mathcal{M} \). Then the entanglement-assisted capacity is given by the expression
\[ C_{ea}(\mathcal{M}, A_E) = \sup_{S, A: \text{Tr} S A F \leq E} \text{ER} (S_A, M). \] (24)

**Proof.** In the proof we use the corresponding result for measurement channels defined by a bounded operator density, obtained in [10].

Let \( S_{AB} \) be the initial entangled state of the system \( AB \). After applying encoding \( \mathcal{E}_A \) in the system \( A \) the state of the composite system is described by the operator
\[ S_{AB}^x = (\mathcal{E}_A^x \otimes \text{Id}_B) S_{AB} \]
with the partial states \( S_A^x = \mathcal{E}_A^x (S_A) \) and \( S_B^x = S_B \).

To establish the inequality \( \leq \) in the formula (24), it is sufficient to prove (see [10] for detail) that
\[ H_{cq} \left( \sum_x \pi_x (\mathcal{M} \otimes \text{Id}_B) S_{AB}^x \right) - \sum_x \pi_x H_{cq} (\mathcal{M} \otimes \text{Id}_B (S_{AB}^x)) \leq \text{ER} (\overline{S}_A, M). \] (25)
Here \( \overline{S}_A = \sum_x \pi_x S_A^x \), and the constraint (13) implies the condition
\[ \text{Tr} \overline{S}_A F \leq E. \] (26)
A result of [10] implies that the relation (25) holds for finite-rank states $S_A$ satisfying the constraint (26). For the proof in the general case we apply approximation by finite-rank states.

Assume first that $\operatorname{Tr} S_A F \leq E' < E$ for a positive $E'$. Let $S_A$ have the spectral decomposition $S_A = \sum_i \lambda_i |\varphi_i\rangle \langle \varphi_i|$. Consider the increasing sequence of projections $P_n = \sum_{i=1}^n |\varphi_i\rangle \langle \varphi_i|$ converging to the unit operator $I_A$, and the sequence of states

$$S_{AB}^x(n) = P_n \otimes I_B S_{AB}^x(n),$$

where $S_{AB}^x(n) = \operatorname{Tr}_A (P_n \otimes I_B S_{AB}^x(n) \otimes I_B)$, $|\phi\rangle$ is a fixed unit vector from $\text{lin}\{\varphi_i\}$, belonging to the domain of $\sqrt{F}$. The partial states of $S_{AB}^x(n)$ are

$$S_A^x(n) = \operatorname{Tr}_B S_{AB}^x(n) = P_n S_A^x(n) + (1 - \operatorname{Tr} P_n S_A^x)|\phi\rangle \langle \phi|, \quad \operatorname{Tr}_A S_{AB}^x(n) = S_B.$$

Then the average state in the system $A$ is equal to

$$S_A(n) = \sum_x \pi_x S_A^x(n) = P_n S_A P_n + (1 - \operatorname{Tr} P_n S_A)|\phi\rangle \langle \phi|.$$

We have $S_A(n) = \sum_{i=1}^n \lambda_i |\varphi_i\rangle \langle \varphi_i| + (1 - \operatorname{Tr} P_n S_A)|\phi\rangle \langle \phi|$, then $\|S_A(n) - S_A\|_1 \to 0$ for $n \to \infty$ and

$$\operatorname{Tr} S_A(n) F = \sum_{i=1}^n \lambda_i \|\sqrt{F} \varphi_i\|^2 + (1 - \operatorname{Tr} P_n S_A)|\phi\rangle \langle F \phi| \leq E' + \varepsilon_n,$$

$\varepsilon_n \to 0$ as $n \to \infty$. Thus, starting from some value of $n$, the density operator $S_A(n)$ satisfies the input constraint (26).

Using the condition (23), similarly to the proof of the coding theorem for measurement of observable in [10] we obtain the inequality (25) for the ensemble $\{\pi_x, S_{AB}^x(n)\}$, which can be written in the following form based on the relative entropy:

$$\sum \pi_x H_{cq} \left( (\mathcal{M} \otimes I_B)(S_{AB}^x(n)) \right) \left\| \sum \pi_x (\mathcal{M} \otimes I_B) S_{AB}^x(n) \right\| \leq \operatorname{ER} (\operatorname{Tr} S_A(n), M).$$

Take the limit $n \to \infty$ in (27). By noting that $\lim_{n \to \infty} H_q(S_A(n)) = H_q(S_A)$, using theorem 2 from [7] (i.e. the equality (10)) in the left-hand side and the lower semicontinuity of the relative entropy in the right-hand side, we obtain (25) for the ensemble $\{\pi_x, S_{AB}^x\}$.
Now consider the case $\text{Tr} \overline{S}_A F = E$. Take a unit vector $|e\rangle \in \text{lin}\{\phi_i\}$, satisfying the condition $\langle e| F e \rangle < E$, and construct the approximation $S_{AB}^x(\varepsilon) = (1 - \varepsilon)S_{AB}^x + \varepsilon|e\rangle\langle e| \otimes S_B$, $0 < \varepsilon < 1$. Then the average state of the system $A$ is $\overline{S}_A(\varepsilon) = (1 - \varepsilon)\overline{S}_A + \varepsilon|e\rangle\langle e|$, and the following condition holds

$$\text{Tr} \overline{S}_A(\varepsilon) F < E.$$ 

Let us repeat previous argument approximating $S_{AB}^x(\varepsilon)$ by the states of the form

$$(1 - \varepsilon)P_n \otimes I_B S_{AB}^x P_n \otimes I_B + |e\rangle\langle e|(S_B - (1 - \varepsilon)\text{Tr}_A P_n \otimes I_B S_{AB}^x)$$

with the partial states $(1 - \varepsilon)S_A^x + (1 - (1 - \varepsilon)\text{Tr}_A S_{AB}^x)|e\rangle\langle e|$ and $S_B$ in the systems $A$ and $B$ correspondingly. We obtain that the inequality (25) holds for $S_{AB}^x(\varepsilon)$, $\overline{S}_A(\varepsilon)$. Since $\lim_{\varepsilon \to 0} H_q(S_A(\varepsilon)) = H_q(\overline{S}_A)$, then, tending $\varepsilon$ to zero we obtain (25) for ensembles satisfying the condition $\text{Tr} \overline{S}_A F = E$. The rest of the proof is similar to the case of observable with a bounded density [10].

To prove the inequality $\geq$ in (24) we consider an arbitrary state $S \in \mathcal{S}(\mathcal{H})$, $S = \sum_{i=1}^{\infty} \lambda_i |\varphi_i\rangle\langle \varphi_i|$, satisfying the input constraint. Apply lemma 1, setting $D = \text{lin}\{\varphi_i\}$ and defining posterior states $\hat{S}(\omega)$ by the relation (7). Then the argument is similar to the proof of proposition 4 from [3], and also theorem 3 from [10]. Theorem 2 is proved.

Of special interest is the case of pure POVM for which there exists a representation (4) of the form

$$\langle \psi|M(A)\psi\rangle = \int_A |\langle a(\omega)|\psi\rangle|^2 \mu(d\omega), \quad \psi \in D.$$ 

(28)

In this case the posterior state (7) is a pure state, not depending on $x$:

$$\hat{S}(\omega) = |e\rangle\langle e|,$$

(29)

where $e \in \mathcal{H}$ is a unit vector. Thus, $H_q(\hat{S}(\omega)) = 0$ and the entropy reduction is equal to

$$\text{ER} (S, M) = H_q(S).$$

(30)

The relation (24) takes the form

$$C_{ea}(\mathcal{M}, \mathcal{A}_E) = \sup_{S_A: \text{Tr} S_A F \leq E} H_q(S).$$

(31)
It is well known that this supremum is attained on the Gibbs state
\[ S_\beta = c(\beta)^{-1} \exp(-\beta F), \quad c(\beta) = \text{Tr} \exp(-\beta F), \quad (32) \]
where \( \beta \) is found from the condition \( \text{Tr} S_\beta F = E \), and it is equal to \( \beta E + c(\beta) \).
Thus theorem \ref{theorem:supremum} implies the following statement.

**Corollary 1.** For arbitrary measurement channel, corresponding to pure POVM,
\[ C_{ea}(\mathcal{M}, \mathcal{A}_E) = \beta E + c(\beta), \]
where \( \beta \) is found from the condition \( \text{Tr} S_\beta F = E \).

Let us illustrate this result by two examples. Let \( \mathcal{H} = L^2(\mathbb{R}) \), \( Q \) be the operator of multiplication by \( x \), \( P = -id/dx \) with the common essential domain \( \mathcal{D} = \mathcal{S}(\mathbb{R}) \) (the space of infinitely differentiable functions rapidly decreasing with all derivatives, see e.g. \cite{8}). The spectral measure \( M \) of the selfadjoint operator \( Q \) can be represented in the form (28):
\[ \langle \psi | M(A) | \psi \rangle = \int_A |\langle x | \psi \rangle|^2 dx, \quad \psi \in \mathcal{S}(\mathbb{R}), \]
where \( \langle x | \psi \rangle = \psi(x), \psi \in \mathcal{S}(\mathbb{R}) \), are the Dirac’s \( \delta \)-functionals. Thus the POVM \( M \) does not have bounded operator density, the result of the paper \cite{10} is not applicable and one should apply the approach of the present paper.
Arbitrary density operator \( S \) in \( L^2(\mathbb{R}) \) is defined by the kernel which is conveniently written in the symbolic form \( \langle x | S | y \rangle \) (for continuous kernels this notation can be understood literally). Consider the channel corresponding to the measurement of observable \( Q \), which maps a density operator \( S \) into the probability density \( \langle x | S | x \rangle \) with respect to the Lebesgue measure on the real line. In quantum optics such a channel describes statistics of homodyne measurement of one mode \( Q, P \) of electromagnetic field \cite{12}. As a constraint operator one usually takes the oscillator energy \( F = (P^2 + Q^2)/2 \). Notice that the condition (23) is fulfilled, as the inequality \( \text{Tr} SF \leq E \) implies
\[ \int x^2 \rho_S(x) \, dx = \text{Tr} SQ^2 \leq 2\text{Tr} SF \leq 2E, \]
and the maximal differential entropy (equal to \( (1/2) \log(2\pi e(2E)) \)) under this constraint is attained on the Gaussian probability density. Substituting this value of supremum, equal to the entropy of the Gibbs state of oscillator with the mean energy \( E \) (see e.g. \cite{12}, \cite{8}) into (31), we obtain
\[ C_{ea}(\mathcal{M}, \mathcal{A}_E) = \left( E + \frac{1}{2} \right) \log \left( E + \frac{1}{2} \right) - \left( E - \frac{1}{2} \right) \log \left( E - \frac{1}{2} \right). \quad (33) \]
On the other hand, the classical capacity of homodyne channel computed in [12], [13] is equal to

\[ C(M_{\text{hom}}, A_E) = \log(2E). \]  

(34)

According to the corollary 1, the relation (33) holds for arbitrary pure measurement channel including heterodyne channel, which maps a density operator \( S \) into probability density \( \langle x, y | S | x, y \rangle \) with respect to the Lebesgue measure on the plane, where \( |x, y\rangle \) are the coherent states of the quantum oscillator [4]. Notice that in this case the bounded operator density exists and results of paper [10] are applicable. The classical capacity of the heterodyne channel computed in [13], [4] is equal to

\[ C(M_{\text{het}}, A_E) = \log \left( E + \frac{1}{2} \right). \]  

(35)

For all \( E > 1/2 \) the inequalities hold

\[ C(M_{\text{het}}, A_E) < C(M_{\text{hom}}, A_E) < C_{\text{ea}}(M, A_E). \]

The graphs of the three capacities are shown on Fig. 1.
References

[1] Bennett C. H., Shor P. W., Smolin J. A., Thapliyal A. V. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. — IEEE Trans. Inform. Theory, 2002, v. 48, No 10, p. 2637–2655.

[2] Berta M., Renes J. M., Wilde M. M. Identifying the information gain of a quantum measurement. arXiv:1301.1594.

[3] Holevo A. S. The classical capacities of quantum channel with input constraint. — Theory Probab. Appl., 2004, v. 48, No 2, p. 243255.

[4] Holevo A. S. Information capacity of quantum observable. — Problems Inform. Transmission, 2012, v. 48, No 1, p. 1–10.

[5] Dobrushin R. L. General formulation of Shannon theorem in information theory. — Russian Math. Surveys, 1959, v. 14, No 6, p. 3–104.

[6] Holevo A. S. Entanglement-breaking channels in infinite dimensions. — Problems Inform. Transmission, 2008, v. 44, No 3, p. 171-184.

[7] Shirokov M. E. Entropy reduction of quantum measurement. — J. Math. Phys., 2011, v. 52, No 5, paper No 052202, 18 p.

[8] Holevo A. S. Quantum systems, channels, information. , 2010, 327.

[9] Barchielli A., Lupieri G. Instruments and mutual entropies in quantum information. — Banach Center Publ., 2006, v. 73, p. 65–80.

[10] Kuznetsova A. A., Holevo A. S. Coding theorems for hybrid channels. — Theory Probab. Appl., 2013, v. 58, No 2, p. 298–324.

[11] Lindblad G. Expectations and entropy inequalities for finite quantum systems. — Comm. Math. Phys., 1974, v. 39, p. 111–119.

[12] Caves C. M., Drummond P. B. Quantum limits of bosonic communication rates. — Rev. Modern Phys., 1994, v. 66, p. 481–538.

[13] Hall M. J. W. Quantum information and correlation bounds. — Phys. Rev. A, 1997, v. 55, No 1, p. 100–113.