

## ON PROPERTIES OF QUANTUM CHANNELS RELATED TO THE CLASSICAL CAPACITY\*

M. E. SHIROKOV†

(Translated by the author)

**Аннотация.** This paper is devoted to further study of the Holevo capacity of infinite dimensional quantum channels. Existence of the unique optimal average state for quantum channel constrained by arbitrary convex set of states is shown. The minimax expression for the Holevo capacity of a constrained channel is obtained.

The  $\chi$ -function and the convex closure of the output entropy of an infinite dimensional quantum channel are considered. It is shown that the  $\chi$ -function of an arbitrary channel is lower semicontinuous on the set of all states and has continuous restrictions to subsets of states with continuous output entropy. The explicit expression for the convex closure of the output entropy of an infinite dimensional quantum channel is obtained and its properties are explored. It is shown that the convex closure of the output entropy coincides with the convex hull of the output entropy on the set of states with finite output entropy and, similarly to the  $\chi$ -function, it has continuous restrictions to subsets of states with continuous output entropy. The applications of the obtained results to the theory of entanglement are considered. The properties of the convex closure of the output entropy make possible to generalize some results related to the additivity problem to the infinite dimensional case.

**Key words.** quantum state, entropy, quantum channel, the Holevo capacity, the  $\chi$ -function, the convex closure of the output entropy of a quantum channel

DOI. 10.1137/S0040585X979XXXX

**1. Introduction.** One of the basic notions of the quantum information theory is the notion of a quantum channel defined as completely positive trace preserving map from the set of states of input quantum system into the set of states of output quantum system. Quantum channels are characterized by a set of capacities defined by type of transmitted information (quantum or classical) and by resources used for this transmission [7], [25].

While major attention in quantum information theory up to now was paid to finite dimensional systems, interest to infinite dimensional systems is increasing. For mathematically rigorous treatment of these systems it is necessary to use specific results from the operator theory in a Hilbert space, measure theory and infinite dimensional convex analysis.

This paper is devoted to study of the Holevo capacity<sup>2</sup> of infinite dimensional quantum channels and related entropic characteristics of quantum channels, following [9], [10], [21], [31].

From mathematical viewpoint the essential features of infinite dimensional quantum channels are noncompactness of state space and discontinuity and unboundedness of the output entropy. As a result for infinite dimensional constrained channel with finite Holevo capacity generally there exist no optimal ensembles, which plays important role in study of finite dimensional channels [29]. In [10] the sufficient condition of existence of an optimal measure (generalized ensemble) for arbitrary constrained

---

\*Received by the editors September 4, 2005. The work was partially supported by the program "Modern problems of theoretical mathematics" of Russian Academy of Sciences and by RFBR (grant N 06-01-00164-a).

<http://www.siam.org/journals/tvp/52-2/9?????.html>

†Steklov Mathematical Institute RAS, Gubkin St. 8, 119991 Moscow, Russia (msh@mi.ras.ru).

<sup>2</sup>This notion is closely related to the classical capacity of quantum channel, see [7].

channel is obtained, but in [10], [32] the examples of constrained channels with no optimal measure are constructed. In section 3 of this paper it is shown that despite possible nonexistence of optimal measure for arbitrary channel constrained by convex set of states there exists the unique state, called output optimal average, which inherits properties of the image of the average state of an optimal ensemble for a finite dimensional constrained channel (proposition 1). The minimax expression for the Holevo capacity is obtained and the alternative definition of the output optimal average state as the minimum point of the lower semicontinuous function on a compact set is given (proposition 2).

In section 4 the notion of the  $\chi$ -function of an infinite dimensional quantum channel is introduced. By using properties of the output optimal average state the inequality for the  $\chi$ -function, derived in [21] for finite dimensional channels, is generalized to the infinite dimensional case (proposition 3). It is shown that the  $\chi$ -function of arbitrary quantum channel is a lower semicontinuous function with natural chain properties (propositions 4–5). The  $\chi$ -function version of Simon’s dominated convergence theorem for quantum entropy is proved (corollary 3).

The another important characteristic of a quantum channel is the convex closure of the output entropy considered in section 5. Since in the finite dimensional case the entropy is a continuous function on a compact set, its convex hull coincides with its convex closure [4] (lower envelope in terms of [12]). The important role of this function in study of finite dimensional channels is justified by its close relation to the  $\chi$ -function: the later is a difference between the output entropy and its convex hull (closure). In the infinite dimensional case the above coincidence do not hold generally and it is natural to consider the convex closure of the output entropy instead of its convex hull. The explicit expression for the convex closure of the output entropy of an infinite dimensional quantum channel is obtained and its properties are explored (propositions 6–8, corollary 4). The main technical problem is noncompactness of the state set, which prevents to use general theory of integral representations on the compact convex sets [2], [12]. The basic instrument of our approach is the criterion of compactness of families of probability measures on the set of quantum states as well as other results from [10]. It is shown that the convex closure of the output entropy coincides with the convex hull of the output entropy on the set of states with finite output entropy. Thus, the representation of the  $\chi$ -function as the difference between the output entropy and its convex closure is valid on this set similarly to the finite dimensional case.

By using the results of the previous sections the following continuity observation is obtained in section 6: the  $\chi$ -function and the convex closure of the output entropy have continuous restrictions to arbitrary set of continuity of the output entropy (theorem 1). This and the observation in [10] imply, in particular, continuity of the  $\chi$ -function for Gaussian channels with constrained mean energy.

Section 7 is devoted to the additivity problem – one of the basic open problems of quantum information theory. The results of the previous sections make possible to prove infinite dimensional versions of the theorems in [21] and [30] concerning equivalence of different additivity properties for given two quantum channels.

The important partial case of the convex closure of the output entropy is a special entanglement measure of a state of a bipartite system called Entanglement of Formation (EoF) [17]. In the finite dimensional case the EoF coincides with the convex closure of the output entropy of a partial trace, which can be considered as a channel from the state space of a bipartite system into the state space of a single subsystem.

In section 8 the arguments for definition of the EoF in the infinite dimensional case as the convex closure of the output entropy of a partial trace are considered. It is shown that this definition is natural and implies such properties of the EoF as convexity, lower semicontinuity on the whole state space and continuity on the set of states with bounded mean energy. It is shown that this definition coincides with the definition proposed in [19] for all states with finite entropy of partial states. The question of their coincidence on the whole state space remains open.

**2. Basic notations.** Let  $\mathcal{H}$  be a separable Hilbert space,  $\mathfrak{B}(\mathcal{H})$  be the algebra of all bounded operators in  $\mathcal{H}$ ,  $\mathfrak{T}(\mathcal{H})$  be the Banach space of all trace class operators with the trace norm  $\|\cdot\|_1$ . In what follows we use the term *state* for positive operator  $\rho$  in  $\mathcal{H}$  with unit trace:  $\rho \geq 0$ ;  $\text{Tr } \rho = 1$ . The algebra  $\mathfrak{B}(\mathcal{H})$  is generally called the algebra of observables of a quantum system while the state  $\rho$  defines the expectation functional  $A \mapsto \text{Tr } \rho A$ ;  $A \in \mathfrak{B}(\mathcal{H})$ , t.i. normal states in terms of the operator algebras theory [2]. The set of all states  $\mathfrak{S}(\mathcal{H})$  is a convex closed subset of  $\mathfrak{T}(\mathcal{H})$ , which is a complete separable metric space with the metric defined by the trace norm.

A finite collection of states  $\{\rho_i\}$  with the corresponding probabilities  $\{\pi_i\}$  is called (finite) *ensemble* and is denoted by  $\{\pi_i, \rho_i\}$ , the state  $\bar{\rho} = \sum_i \pi_i \rho_i$  is called *the average state* of this ensemble. In [10] the notion of *generalized ensemble* as Borel probability measure  $\pi$  on  $\mathfrak{S}(\mathcal{H})$  is introduced. *The average state* of a generalized ensemble  $\pi$  is the state (also called barycenter of the measure  $\pi$ ), defined by the Bochner integral

$$\bar{\rho}(\pi) = \int_{\mathfrak{S}(\mathcal{H})} \rho \pi(d\rho).$$

Conventional ensembles correspond to measures with finite support.

A convex combination of ensembles is defined as a convex combination of corresponding probability measures. In particular, for arbitrary set of ensembles  $\{\{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)}\}_{k=1}^m$  and probability distribution  $\{\lambda_k\}_{k=1}^m$  the convex combination  $\sum_{k=1}^m \lambda_k \{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)}$  of the above ensembles is the ensemble, consisting of  $\sum_{k=1}^m n(k)$  states  $\{\rho_i^k\}_{k,i}$  with the corresponding probabilities  $\{\lambda_k \pi_i^k\}_{k,i}$ .

Let  $\mathcal{P}$  be the convex set of all probability measures on  $\mathfrak{S}(\mathcal{H})$ , endowed with the topology of weak convergence [1]. In [10] it is noted that the map  $\mathcal{P} \ni \pi \mapsto \bar{\rho}(\pi)$  is continuous in the above topology. The subset of  $\mathcal{P}$ , consisting of all measures  $\pi$  with the barycenter  $\bar{\rho}(\pi)$  in  $\mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$ , will be denoted by  $\mathcal{P}_{\mathcal{A}}$ .

Let  $A$  and  $B$  be positive operators in  $\mathfrak{T}(\mathcal{H})$ . The von Neumann entropy of the operator  $A$  and the relative entropy of the operators  $A$  and  $B$  are defined correspondingly by the expressions

$$H(A) = -\sum_i \langle i | A \log A | i \rangle \quad \text{and} \quad H(A||B) = \sum_i \langle i | A \log A - A \log B + B - A | i \rangle,$$

where  $\{|i\rangle\}$  is the basis of eigenvectors of the operator  $A$  (see details in [22], [34]). The entropy and the relative entropy are lower semicontinuous functions of their arguments taking values in  $[0; +\infty]$ , the first function is concave while the second one is convex [22], [34]. The following inequality

$$(1) \quad H(\rho||\sigma) \geq \frac{1}{2} \|\rho - \sigma\|_1^2,$$

holds for arbitrary states  $\rho$  and  $\sigma$  in  $\mathfrak{S}(\mathcal{H})$  [27].

The relative entropy  $H(\rho\|\sigma)$  of states  $\rho$  and  $\sigma$  can be considered as a measure of divergence of these states, its classical analog is called Kulback-Leibler distance. Despite the fact that this measure is not a metric (it is nonsymmetric and do not satisfy the triangle inequality), it is possible to introduce the notion of convergence of a sequence of states  $\{\rho_n\}$  to a particular state  $\rho_*$ , defined by the condition  $\lim_{n\rightarrow+\infty} H(\rho_n\|\rho_*) = 0$ . In the classical case the topology on the state space, related with this convergence, is explored in [20], where it is called strong information topology. In quantum information theory this convergence, which will be called *H-convergence*, also plays important role (see. [11, proposition 2]). It follows from inequality (1) that

$$\left\{ \lim_{n\rightarrow+\infty} H(\rho_n\|\rho_*) = 0 \right\} \Leftrightarrow \left\{ \lim_{n\rightarrow+\infty} \rho_n = \rho_* \right\}.$$

We will use Donald's identity [15], [27]

$$(2) \quad \sum_{i=1}^n \pi_i H(\rho_i\|\bar{\rho}) = \sum_{i=1}^n \pi_i H(\rho_i\|\bar{\rho}) + H(\bar{\rho}\|\bar{\rho}),$$

which holds for arbitrary ensemble  $\{\pi_i, \rho_i\}_{i=1}^n$  with the average state  $\bar{\rho}$  and arbitrary state  $\hat{\rho}$ .

Let  $\mathcal{H}$ ,  $\mathcal{H}'$  be a pair of separable Hilbert spaces, which will be called correspondingly input and output spaces. Channel  $\Phi$  is a linear positive trace preserving map from  $\mathfrak{T}(\mathcal{H})$  into  $\mathfrak{T}(\mathcal{H}')$  such that the dual map  $\Phi^*: \mathfrak{B}(\mathcal{H}') \mapsto \mathfrak{B}(\mathcal{H})$  (which exists since  $\Phi$  is bounded) is completely positive, see. [8, ch. 3, p. 1]. In particular, channel maps *input* states in  $\mathfrak{S}(\mathcal{H})$  into *output* states in  $\mathfrak{S}(\mathcal{H}')$ .

The important characteristic of a channel  $\Phi$  is its output entropy  $H_\Phi(\rho) = H(\Phi(\rho))$  – concave lower semicontinuous nonnegative function on the set of input states  $\mathfrak{S}(\mathcal{H})$ .

Let  $\mathcal{A}$  be an arbitrary subset of  $\mathfrak{S}(\mathcal{H})$ . Consider the constraint on input ensemble  $\{\pi_i, \rho_i\}$ , defined by the inclusion  $\bar{\rho} \in \mathcal{A}$ . A channel  $\Phi$  with this constraint is called  *$\mathcal{A}$ -constrained* channel. The *Holevo capacity* of the  $\mathcal{A}$ -constrained channel  $\Phi$  is defined as follows [9], [10]:

$$(3) \quad \bar{C}(\Phi, \mathcal{A}) = \sup_{\bar{\rho} \in \mathcal{A}} \chi_\Phi(\{\pi_i, \rho_i\}),$$

where

$$(4) \quad \chi_\Phi(\{\pi_i, \rho_i\}) = \sum_i \pi_i H(\Phi(\rho_i)\|\Phi(\bar{\rho})).$$

In [10] it is shown that the Holevo capacity of the  $\mathcal{A}$ -constrained channel  $\Phi$  can be also defined by the expression

$$(5) \quad \bar{C}(\Phi, \mathcal{A}) = \sup_{\pi \in \mathcal{P}_\mathcal{A}} \int_{\mathfrak{S}(\mathcal{H})} H(\Phi(\rho)\|\Phi(\bar{\rho}(\pi))) \pi(d\rho),$$

which means coincidence of the supremum over all measures in  $\mathcal{P}_\mathcal{A}$  with the supremum over finitely supported measures in  $\mathcal{P}_\mathcal{A}$ .

**3. The optimal average state.** It is well known that for arbitrary finite dimensional channel  $\Phi$  and arbitrary closed set  $\mathcal{A}$  there exists optimal ensemble  $\{\pi_i, \rho_i\}$ , at which the supremum in definition (3) of the Holevo capacity is achieved

[18], [29]. The image of the average state of an optimal ensemble plays important role in study of finite dimensional channels [21].

For infinite dimensional quantum channels generally there exist no optimal finite ensembles. Thus it is natural to introduce the notion of generalized optimal ensemble as a probability measure on the set of all states, at which the supremum in definition (5) of the Holevo capacity is achieved. In [10] the sufficient condition of existence of optimal measure for arbitrary constrained channel is obtained, but examples in [10], [32] shows that optimal measure does not exist in some cases.

The aim of this section is to show that despite possible nonexistence of an optimal measure there exists the unique state, inheriting basic properties of the image of the average state of optimal ensemble for finite dimensional channel. This state is naturally called the output optimal average for constrained channel. If optimal measure exists then the image of its average state (barycenter) coincides with the above output optimal average. By using properties of this state it is possible to generalize some results in [21] to the case of infinite dimensional channels.

LEMMA 1. *Let  $\{\{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)}\}_{k=1}^m$  be a finite set of ensembles and  $\{\lambda_k\}_{k=1}^m$  be a probability distribution. Then*

$$\chi_{\Phi} \left( \sum_{k=1}^m \lambda_k \{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)} \right) = \sum_{k=1}^m \lambda_k \chi_{\Phi} \left( \{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)} \right) + \chi_{\Phi} \left( \{\lambda_k, \bar{\rho}_k\}_{k=1}^m \right),$$

where  $\bar{\rho}_k$  is the average state of the ensemble  $\{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)}$ ,  $k = 1, \dots, m$ .

If  $m = 2$  then for arbitrary  $\lambda \in [0, 1]$  the following inequality holds

$$\begin{aligned} \chi_{\Phi} \left( \lambda \{\pi_i^1, \rho_i^1\}_{i=1}^{n(1)} + (1 - \lambda) \{\pi_i^2, \rho_i^2\}_{i=1}^{n(2)} \right) &\geq \lambda \chi_{\Phi} \left( \{\pi_i^1, \rho_i^1\}_{i=1}^{n(1)} \right) \\ &+ (1 - \lambda) \chi_{\Phi} \left( \{\pi_i^2, \rho_i^2\}_{i=1}^{n(2)} \right) + \frac{\lambda(1 - \lambda)}{2} \|\Phi(\bar{\rho}_2) - \Phi(\bar{\rho}_1)\|_1^2. \end{aligned}$$

*Proof.* Let  $\bar{\rho} = \sum_{k=1}^m \lambda_k \bar{\rho}_k$  be the average state of the ensemble  $\sum_{k=1}^m \lambda_k \{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)}$ . By definition

$$\chi_{\Phi} \left( \sum_{k=1}^m \lambda_k \{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)} \right) = \sum_{k=1}^m \lambda_k \sum_{i=1}^{n(k)} \pi_i^k H(\Phi(\rho_i^k) \| \Phi(\bar{\rho})).$$

By applying Donald's identity (2) to each inner sum in the right side of this expression we obtain the main identity of the lemma.

To prove inequality for  $m = 2$  it is sufficient to use inequality (1) for estimation of the last term in the right side of the main identity of the lemma:

$$\begin{aligned} &\lambda H(\Phi(\bar{\rho}_1) \| \Phi(\lambda \bar{\rho}_1 + (1 - \lambda) \bar{\rho}_2)) + (1 - \lambda) H(\Phi(\bar{\rho}_2) \| \Phi(\lambda \bar{\rho}_1 + (1 - \lambda) \bar{\rho}_2)) \\ &\geq \frac{1}{2} \lambda \|(1 - \lambda) \Phi(\bar{\rho}_2 - \bar{\rho}_1)\|_1^2 + \frac{1}{2} (1 - \lambda) \|\lambda \Phi(\bar{\rho}_2 - \bar{\rho}_1)\|_1^2 \\ &= \frac{1}{2} \lambda(1 - \lambda) \|\Phi(\bar{\rho}_2) - \Phi(\bar{\rho}_1)\|_1^2. \end{aligned}$$

Lemma 1 is proved.

Despite possible nonexistence of optimal ensemble for the  $\mathcal{A}$ -constrained channel  $\Phi$ , the definition of the Holevo capacity implies existence of the sequence with the following properties.

DEFINITION 1. A sequence of ensembles  $\{\{\pi_i^k, \rho_i^k\}\}_k$  such that  $\bar{\rho}_k = \sum_i \pi_i^k \rho_i^k \in \mathcal{A}$  for all  $k$  and  $\lim_{k \rightarrow +\infty} \chi_\Phi(\{\pi_i^k, \rho_i^k\}) = \bar{C}(\Phi, \mathcal{A})$  is called *approximating sequence* for the  $\mathcal{A}$ -constrained channel  $\Phi$ .

A state is called *input optimal average* for the  $\mathcal{A}$ -constrained channel  $\Phi$ , if it is a limit of the sequence of the average states of some approximating sequence of ensembles for the  $\mathcal{A}$ -constrained channel  $\Phi$ .

This definition guarantees neither existence nor uniqueness of input optimal average (the examples of channels with no input optimal average are considered in [3]). Existence of at least one input optimal average is a necessary condition of existence of optimal measure for the  $\mathcal{A}$ -constrained channel  $\Phi$ , which also becomes a sufficient condition if some additional requirements hold. This input optimal average coincides with the barycenter of the optimal measure (see details in [10]).

Despite possible nonexistence of limit points of the sequence of the average states of approximating sequence of ensembles for the  $\mathcal{A}$ -constrained channel, finiteness of its Holevo capacity guarantees convergence of the sequence of the images of these averages.

The following proposition is a generalization of proposition 1 in [31] to the case of noncompact set  $\mathcal{A}$ .

PROPOSITION 1. Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  be an arbitrary channel and  $\mathcal{A}$  be a convex subset of  $\mathfrak{S}(\mathcal{H})$  such that  $\bar{C}(\Phi, \mathcal{A}) < +\infty$ . There exists the unique state  $\Omega(\Phi, \mathcal{A})$  in  $\mathfrak{S}(\mathcal{H}')$  such that  $\sum_j \mu_j H(\Phi(\sigma_j) \| \Omega(\Phi, \mathcal{A})) \leq \bar{C}(\Phi, \mathcal{A})$  for any ensemble  $\{\mu_j, \sigma_j\}$  with the average  $\bar{\sigma} \in \mathcal{A}$ .

The state  $\Omega(\Phi, \mathcal{A})$  lies in  $\overline{\Phi(\mathcal{A})}$ . For arbitrary approximating sequence of ensembles  $\{\{\pi_i^k, \rho_i^k\}\}_k$  for the  $\mathcal{A}$ -constrained channel  $\Phi$  with the corresponding sequence of the average states  $\{\bar{\rho}_k\}_k$  there exists

$$H\text{-}\lim_{k \rightarrow +\infty} \Phi(\bar{\rho}_k) = \Omega(\Phi, \mathcal{A}).$$

*Proof.* Show first that for arbitrary approximating sequence of ensembles  $\{\Sigma_k = \{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)}\}$  for the  $\mathcal{A}$ -constrained channel  $\Phi$  the sequence  $\{\Phi(\bar{\rho}_k)\}$  converges to some state in  $\mathfrak{S}(\mathcal{H}')$ . By definition of approximating sequence for arbitrary  $\varepsilon > 0$  there exists  $N_\varepsilon$  such that  $\chi_\Phi(\Sigma_k) > \bar{C}(\Phi, \mathcal{A}) - \varepsilon$  for any  $k \geq N_\varepsilon$ . By lemma 1 (with  $m = 2$  and  $\lambda = \frac{1}{2}$ ) we have

$$\begin{aligned} \bar{C}(\Phi, \mathcal{A}) - \varepsilon &\leq \frac{1}{2} \chi_\Phi(\Sigma_{k_1}) + \frac{1}{2} \chi_\Phi(\Sigma_{k_2}) \leq \chi_\Phi\left(\frac{1}{2} \Sigma_{k_1} + \frac{1}{2} \Sigma_{k_2}\right) \\ &\quad - \frac{1}{8} \|\Phi(\bar{\rho}_{k_2}) - \Phi(\bar{\rho}_{k_1})\|_1^2 \leq \bar{C}(\Phi, \mathcal{A}) - \frac{1}{8} \|\Phi(\bar{\rho}_{k_2}) - \Phi(\bar{\rho}_{k_1})\|_1^2, \end{aligned}$$

for any  $k_1 \geq N_\varepsilon$  and  $k_2 \geq N_\varepsilon$ . Hence,  $\|\Phi(\bar{\rho}_{k_2}) - \Phi(\bar{\rho}_{k_1})\|_1 < \sqrt{8\varepsilon}$ . Thus the sequence  $\{\Phi(\bar{\rho}_k)\}$  is a Cauchy sequence and hence it converges to some state  $\omega$  in  $\mathfrak{S}(\mathcal{H}')$ .

Let  $\{\mu_j, \sigma_j\}_{j=1}^m$  be arbitrary ensemble of  $m$  state with the average  $\bar{\sigma} \in \mathcal{A}$ . Consider the family of ensembles

$$\Sigma_k^\lambda = (1 - \lambda) \{\pi_i^k, \rho_i^k\}_{i=1}^{n(k)} + \lambda \{\mu_j, \sigma_j\}_{j=1}^m, \quad \lambda \in [0, 1], \quad k \in \mathbf{N},$$

with the average states  $\bar{\rho}_k^\lambda = (1 - \lambda) \bar{\rho}_k + \lambda \bar{\sigma}$ . By convexity of the set  $\mathcal{A}$  we have  $\bar{\rho}_k^\lambda \in \mathcal{A}$  for all  $\lambda \in [0, 1]$  and  $k \in \mathbf{N}$  and the above observation implies

$$(6) \quad \lim_{k \rightarrow +\infty} \Phi(\bar{\rho}_k^\lambda) = (1 - \lambda) \omega + \lambda \Phi(\bar{\sigma}).$$

By definition

$$(7) \quad \chi_{\Phi}(\Sigma_k^{\lambda}) = (1 - \lambda) \sum_{i=1}^{n(k)} \pi_i^k H(\Phi(\rho_i^k) \parallel \Phi(\bar{\rho}_k^{\lambda})) + \lambda \sum_{j=1}^m \mu_j H(\Phi(\sigma_j) \parallel \Phi(\bar{\rho}_k^{\lambda})).$$

By the condition  $\bar{C}(\Phi, \mathcal{A}) < +\infty$  the both sums in the right side of this expression are finite. Applying Donald's identity (2) to the first sum we obtain

$$\sum_{i=1}^{n(k)} \pi_i^k H(\Phi(\rho_i^k) \parallel \Phi(\bar{\rho}_k^{\lambda})) = \chi_{\Phi}(\Sigma_k^0) + H(\Phi(\bar{\rho}_k) \parallel \Phi(\bar{\rho}_k^{\lambda})).$$

Substitution of this expression to (7) leads to

$$\begin{aligned} \chi_{\Phi}(\Sigma_k^{\lambda}) &= \chi_{\Phi}(\Sigma_k^0) + (1 - \lambda) H(\Phi(\bar{\rho}_k) \parallel \Phi(\bar{\rho}_k^{\lambda})) \\ &\quad + \lambda \left( \sum_{j=1}^m \mu_j H(\Phi(\sigma_j) \parallel \Phi(\bar{\rho}_k^{\lambda})) - \chi_{\Phi}(\Sigma_k^0) \right). \end{aligned}$$

Thus by nonnegativity of the relative entropy we obtain

$$(8) \quad \sum_{j=1}^m \mu_j H(\Phi(\sigma_j) \parallel \Phi(\bar{\rho}_k^{\lambda})) \leq \lambda^{-1} (\chi_{\Phi}(\Sigma_k^{\lambda}) - \chi_{\Phi}(\Sigma_k^0)) + \chi_{\Phi}(\Sigma_k^0)$$

for  $\lambda \neq 0$ . The definition of approximating sequence implies

$$(9) \quad \lim_{k \rightarrow +\infty} \chi_{\Phi}(\Sigma_k^0) = \bar{C}(\Phi, \mathcal{A}) \geq \chi_{\Phi}(\Sigma_k^{\lambda})$$

for all  $k$  and all  $\lambda \in [0, 1]$ . Thus

$$(10) \quad \liminf_{\lambda \rightarrow +0} \liminf_{k \rightarrow +\infty} \lambda^{-1} (\chi_{\Phi}(\Sigma_k^{\lambda}) - \chi_{\Phi}(\Sigma_k^0)) \leq 0.$$

By lower semicontinuity of the relative entropy it follows from (6), (8), (9) and (10) that

$$\sum_{j=1}^m \mu_j H(\Phi(\sigma_j) \parallel \omega) \leq \liminf_{\lambda \rightarrow +0} \liminf_{k \rightarrow +\infty} \sum_{j=1}^m \mu_j H(\Phi(\sigma_j) \parallel \Phi(\bar{\rho}_k^{\lambda})) \leq \bar{C}(\Phi, \mathcal{A}).$$

Thus it is proved that

$$(11) \quad \sum_j \mu_j H(\Phi(\sigma_j) \parallel \omega) \leq \bar{C}(\Phi, \mathcal{A})$$

for arbitrary ensemble  $\{\mu_j, \sigma_j\}$  with the average state  $\bar{\sigma} \in \mathcal{A}$ .

Let  $\{\{\mu_i^k, \sigma_i^k\}\}_k$  be arbitrary approximating sequence of ensembles for the  $\mathcal{A}$ -constrained channel  $\Phi$  with the corresponding sequence of the average states  $\{\bar{\sigma}_k\}_k$ . Property (11) implies

$$\sum_i \mu_i^k H(\Phi(\sigma_i^k) \parallel \omega) \leq \bar{C}(\Phi, \mathcal{A}) \quad \forall k.$$

Applying identity (2) we obtain

$$(12) \quad \sum_i \mu_i^k H(\Phi(\sigma_i^k) \parallel \omega) = \sum_i \mu_i^k H(\Phi(\sigma_i^k) \parallel \Phi(\bar{\sigma}_k)) + H(\Phi(\bar{\sigma}_k) \parallel \omega).$$

The last two expressions imply

$$H(\Phi(\bar{\sigma}_k) \parallel \omega) \leq \bar{C}(\Phi, \mathcal{A}) - \sum_i \mu_i^k H(\Phi(\sigma_i^k) \parallel \Phi(\bar{\sigma}_k)).$$

By approximating property of the sequence  $\{\{\mu_i^k, \sigma_i^k\}\}_k$  the right side of this inequality tends to zero as  $k \rightarrow +\infty$ , hence there exists  $\lim_{k \rightarrow +\infty} H(\Phi(\bar{\sigma}_k) \parallel \omega) = 0$ , which implies convergence of the sequence  $\{\bar{\sigma}_k\}$  to the state  $\omega$ . Thus the state  $\omega$  does not depend on the choice of approximating sequence and hence it is defined only by the channel  $\Phi$  and the set  $\mathcal{A}$ . Denote this state by  $\Omega(\Phi, \mathcal{A})$ . The above observation also implies that  $\omega = \Omega(\Phi, \mathcal{A})$  is the unique state such that property (11) holds. Proposition 1 is proved.

Proposition 1 shows in particular that the set of input optimal averages for the  $\mathcal{A}$ -constrained channel  $\Phi$  is either empty or mapped by the channel  $\Phi$  to the single state.

**COROLLARY 1.** *Let  $\mathcal{A}$  be a convex subset of  $\mathfrak{S}(\mathcal{H})$ . If there exists input optimal average  $\rho_*$  for the  $\mathcal{A}$ -constrained channel  $\Phi$  then  $\Phi(\rho_*) = \Omega(\Phi, \mathcal{A})$ .*

If the set  $\mathcal{A}$  is compact then there exists at least one input optimal average state.

This corollary justifies the following definition.

**DEFINITION 2.** The state  $\Omega(\Phi, \mathcal{A})$  is called the *output optimal average* for the  $\mathcal{A}$ -constrained channel  $\Phi$ .

Note that there exist examples of the  $\mathcal{A}$ -constrained channels  $\Phi$  with finite Holevo capacity with no input optimal averages while the output optimal average  $\Omega(\Phi, \mathcal{A})$  is explicitly determined and plays important role in analysis of these channels (see examples in [3], [32]).

There exists the another approach to the definition of the state  $\Omega(\Phi, \mathcal{A})$ . In [11] it is shown (corollary 6) that finiteness of the Holevo capacity of the  $\mathcal{A}$ -constrained channel  $\Phi$  implies compactness of the set  $\bar{\Phi}(\mathcal{A})$ . For any ensemble  $\{\mu_j, \sigma_j\}$  with the average  $\bar{\sigma} \in \mathcal{A}$  consider lower semicontinuous function  $F_{\{\mu_j, \sigma_j\}}(\omega) = \sum_j \mu_j H(\Phi(\sigma_j) \parallel \omega)$  on the set  $\bar{\Phi}(\mathcal{A})$ . The function  $F(\omega) = \sup_{\sum_j \mu_j \sigma_j \in \mathcal{A}} F_{\{\mu_j, \sigma_j\}}(\omega)$  is also lower semicontinuous and hence it achieves its minimum on the compact set  $\bar{\Phi}(\mathcal{A})$ . The following proposition shows in particular that the state  $\Omega(\Phi, \mathcal{A})$  can be defined as the unique minimum point of the function  $F(\omega)$  on the set  $\bar{\Phi}(\mathcal{A})$ .

**PROPOSITION 2.** *Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  be an arbitrary channel and  $\mathcal{A}$  be a convex subset of  $\mathfrak{S}(\mathcal{H})$ . The Holevo capacity of the  $\mathcal{A}$ -constrained channel  $\Phi$  is defined by the expression*

$$\bar{C}(\Phi, \mathcal{A}) = \inf_{\omega \in \bar{\Phi}(\mathcal{A})} \left( \sup_{\sum_j \mu_j \sigma_j \in \mathcal{A}} \sum_j \mu_j H(\Phi(\sigma_j) \parallel \omega) \right).$$

*If  $\bar{C}(\Phi, \mathcal{A}) < +\infty$  then  $\Omega(\Phi, \mathcal{A})$  is the unique state at which the infimum in the right side of this expression is achieved.*

*Proof.* Let  $\bar{C}(\Phi, \mathcal{A}) < +\infty$ . Show first that

$$(13) \quad \sup_{\sum_j \mu_j \sigma_j \in \mathcal{A}} \sum_j \mu_j H(\Phi(\sigma_j) \parallel \Omega(\Phi, \mathcal{A})) = \bar{C}(\Phi, \mathcal{A}).$$



Proposition 1 implies inequality “ $\leq$ ” in (13). To prove the converse inequality consider arbitrary approximating sequence  $\{\{\pi_i^k, \rho_i^k\}\}_k$ . By using Donald’s identity (2) we obtain

$$\sum_i \pi_i^k H(\Phi(\rho_i^k) \| \Omega(\Phi, \mathcal{A})) = \sum_i \pi_i^k H(\Phi(\rho_i^k) \| \Phi(\bar{\rho}_k)) + H(\Phi(\bar{\rho}_k) \| \Omega(\Phi, \mathcal{A}))$$

for all  $k$ . By approximating property of the sequence  $\{\{\pi_i^k, \rho_i^k\}\}_k$  the first term in the right side tends to  $\bar{C}(\Phi, \mathcal{A})$  as  $k \rightarrow +\infty$  while the second one is nonnegative. This proves inequality “ $\geq$ ” and hence equality in (13).

Let  $\omega_*$  be a minimal point of the function  $F(\omega)$ . By equality (13) we have

$$\sup_{\sum_j \mu_j \sigma_j \in \mathcal{A}} \sum_j \mu_j H(\Phi(\sigma_j) \| \omega_*) = F(\omega_*) \leq F(\Omega(\Phi, \mathcal{A})) = \bar{C}(\Phi, \mathcal{A}).$$

Proposition 1 implies  $\omega_* = \Omega(\Phi, \mathcal{A})$ .

If  $\bar{C}(\Phi, \mathcal{A}) = +\infty$  then the right side of the expression in proposition 2 is equal to  $+\infty$ . Indeed, if there exists state  $\omega$  in  $\mathfrak{S}(\mathcal{H}')$  such that  $\sup_{\sum_j \mu_j \sigma_j \in \mathcal{A}} \sum_j \mu_j H(\Phi(\sigma_j) \| \omega) < +\infty$  then equality (12), valid for arbitrary approximating sequence of ensembles  $\{\{\mu_i^k, \sigma_i^k\}\}_k$ , implies  $\bar{C}(\Phi, \mathcal{A}) < +\infty$ . Proposition 2 is proved. ▀

*Remark 1.* The assumption of convexity of the set  $\mathcal{A}$  in propositions 1, 2 and corollary 1 is essential. Consider the noiseless channel  $\Phi = \text{Id}$  and the set  $\mathcal{A}$  consisting of two states  $\rho_1$  and  $\rho_2$  such that  $H(\rho_1) = H(\rho_2) < +\infty$ . In this case  $\bar{C}(\Phi, \mathcal{A}) = H(\rho_1) = H(\rho_2)$ , the states  $\rho_1$  and  $\rho_2$  are input optimal averages in the sense of definition 1 with different images  $\Phi(\rho_1) = \rho_1$  and  $\Phi(\rho_2) = \rho_2$ .

**4.  $\chi$ -function.** Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  be an arbitrary channel. Consider the function  $\chi_\Phi$  on the set  $\mathfrak{S}(\mathcal{H})$ , which takes value  $\bar{C}(\Phi, \{\rho\})$  at state  $\rho$ . By using definitions of the Holevo capacity (3) and (5), we obtain

$$(14) \quad \chi_\Phi(\rho) = \sup_{\sum_i \pi_i \rho_i = \rho} \sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho)) = \sup_{\pi \in \mathcal{P}_{\{\rho\}}} \int_{\mathfrak{S}(\mathcal{H})} H(\Phi(\sigma) \| \Phi(\rho)) \pi(d\sigma).$$

In the finite dimensional case  $\chi_\Phi$  is a continuous concave nonnegative function on the set of input states  $\mathfrak{S}(\mathcal{H})$ , used in study of the classical capacity of quantum channels, in particular, of the additivity problem [21]. In this section the properties of the  $\chi$ -function of arbitrary infinite dimensional channel  $\Phi$  are considered.

In [10] it is shown that if  $H_\Phi(\rho) < +\infty$  then the supremum in the last expression in (14) is achieved at some measure supported by the set of pure states.

**DEFINITION 3.** A measure  $\pi_0$  with the barycenter  $\rho_0$  supported by the set of pure states such that

$$\chi_\Phi(\rho_0) = \int_{\mathfrak{S}(\mathcal{H})} H(\Phi(\sigma) \| \Phi(\rho_0)) \pi_0(d\sigma),$$

is called  $\chi_\Phi$ -optimal measure for the state  $\rho_0$ .

Note that  $\chi_\Phi(\rho) < +\infty$  does not imply  $H_\Phi(\rho) < +\infty$ . Indeed, it is easy to construct the channel  $\Phi$  from finite dimensional into infinite dimensional spaces such that  $H_\Phi(\rho) = +\infty$  for any state  $\rho$  in  $\mathfrak{S}(\mathcal{H})$ .<sup>3</sup> By monotonicity property of the relative

<sup>3</sup>For example the channel  $\Phi: \mathfrak{S}(\mathcal{H}) \ni A \mapsto \frac{1}{2} A \oplus \frac{1}{2} \sigma \text{Tr} A \in \mathfrak{S}(\mathcal{H}')$ , where  $\sigma$  is a fixed state with infinite entropy.

entropy [23] we have

$$\sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho)) \leq \sum_i \pi_i H(\rho_i \| \rho) \leq \log \dim \mathcal{H} < +\infty$$

for any ensemble  $\{\pi_i, \rho_i\}$  and hence  $\chi_\Phi(\rho) \leq \log \dim \mathcal{H} < +\infty$  for any state  $\rho$  in  $\mathfrak{S}(\mathcal{H})$ .

The definition of the Holevo capacity of the  $\mathcal{A}$ -constrained channel  $\Phi$  implies  $\overline{C}(\Phi, \mathcal{A}) = \sup_{\rho \in \mathcal{A}} \chi_\Phi(\rho)$ . In [10] it is shown that the question of attainability of the above supremum is closely related with the question of existence of optimal measure for the  $\mathcal{A}$ -constrained channel  $\Phi$  and due to examples in [32] it can have negative answer.

The results of the previous section make possible to prove the important inequality, determining behavior of the  $\chi$ -function on any convex subset of states, which is proved in [21] in the finite dimensional case. This inequality plays the essential role in the proof of theorem 2 in section 7.

**PROPOSITION 3.** *Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  be an arbitrary channel and  $\mathcal{A}$  be a convex subset of  $\mathfrak{S}(\mathcal{H})$ . Then for any state  $\rho$  in  $\mathcal{A}$  the following inequality holds*

$$\chi_\Phi(\rho) \leq \overline{C}(\Phi, \mathcal{A}) - H(\Phi(\rho) \| \Omega(\Phi, \mathcal{A})) \leq \overline{C}(\Phi, \mathcal{A}) - \frac{1}{2} \|\Phi(\rho) - \Omega(\Phi, \mathcal{A})\|_1^2.$$

*Proof.* Let  $\overline{C}(\Phi, \mathcal{A}) < +\infty$  and  $\{\pi_i, \rho_i\}$  be arbitrary ensemble such that  $\sum_i \pi_i \rho_i = \rho$ . By proposition 1

$$\sum_i \pi_i H(\Phi(\rho_i) \| \Omega(\Phi, \mathcal{A})) \leq \overline{C}(\Phi, \mathcal{A}).$$

This inequality, Donald's identity

$$\sum_i \pi_i H(\Phi(\rho_i) \| \Omega(\Phi, \mathcal{A})) = \sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho)) + H(\Phi(\rho) \| \Omega(\Phi, \mathcal{A}))$$

and definition of the  $\chi$ -function (14) imply the first inequality of the proposition. The second one follows from inequality (1). Proposition 3 is proved.

For arbitrary state  $\rho$  with finite output entropy the  $\chi$ -function has the following representation:

$$(15) \quad \chi_\Phi(\rho) = H_\Phi(\rho) - \text{co } H_\Phi(\rho),$$

where

$$(16) \quad \text{co } H_\Phi(\rho) = \inf_{\sum_i \pi_i \rho_i = \rho} \sum_i \pi_i H_\Phi(\rho_i).$$

is the convex hull of the output entropy (see. appendix A).

In the finite dimensional case the output entropy  $H_\Phi$  and its convex hull  $\text{co } H_\Phi$  are continuous functions on the set  $\mathfrak{S}(\mathcal{H})$ , the first function is concave while the second one is convex and representation (15) holds for all states. Hence in this case the function  $\chi_\Phi$  is continuous and concave on the set  $\mathfrak{S}(\mathcal{H})$ .

In the infinite dimensional case the output entropy  $H_\Phi$  is only lower semicontinuous and hence the function  $\chi_\Phi$  is not continuous even in the case of the noiseless channel  $\Phi$  for which  $\chi_\Phi = H_\Phi$ . The following proposition shows that the function  $\chi_\Phi$  of arbitrary channel  $\Phi$  has properties similar to the properties of the output entropy  $H_\Phi$ .

PROPOSITION 4. For arbitrary channel  $\Phi$  the function  $\chi_\Phi$  is a nonnegative lower semicontinuous concave function on the set  $\mathfrak{S}(\mathcal{H})$  and

$$(17) \quad \chi_\Phi(\rho) - \sum_{i=1}^n \pi_i \chi_\Phi(\rho_i) \geq \sum_{i=1}^n \pi_i H(\Phi(\rho_i) \| \Phi(\rho))$$

for any ensemble of states  $\{\pi_i, \rho_i\}_{i=1}^n$  with the average  $\rho$ .

Inequality (17) can be considered as a generalization to the case of the  $\chi$ -function of the well known identity for quantum entropy

$$H(\rho) - \sum_{i=1}^n \pi_i H(\rho_i) = \sum_{i=1}^n \pi_i H(\rho_i \| \rho).$$

*Proof.* Nonnegativity of the  $\chi$ -function is obvious. Lower semicontinuity is proved in [31] (proposition 3). To show concavity of the  $\chi$ -function it is sufficient to prove inequality (17). Let  $\varepsilon > 0$ . By the definition of the  $\chi$ -function for each  $i = 1, \dots, n$  there exists ensemble  $\{\mu_j^i, \sigma_j^i\}_{j=1}^{m(i)}$  with the average state  $\rho_i$  such that  $\chi_\Phi(\{\mu_j^i, \sigma_j^i\}) > \chi_\Phi(\rho_i) - \varepsilon$ . Since the average state of the ensemble  $\sum_{i=1}^n \pi_i \{\mu_j^i, \sigma_j^i\}$  coincides with  $\rho$ , by using lemma 1 we obtain

$$\begin{aligned} \chi_\Phi(\rho) &\geq \chi_\Phi\left(\sum_{i=1}^n \pi_i \{\mu_j^i, \sigma_j^i\}\right) = \sum_{i=1}^n \pi_i \chi_\Phi(\{\mu_j^i, \sigma_j^i\}) + \sum_{i=1}^n \pi_i H(\Phi(\rho_i) \| \Phi(\rho)) \\ &\geq \sum_{i=1}^n \pi_i \chi_\Phi(\rho_i) + \sum_{i=1}^n \pi_i H(\Phi(\rho_i) \| \Phi(\rho)) - \varepsilon. \end{aligned}$$

Since  $\varepsilon$  is arbitrary this implies inequality (17). Proposition 4 is proved.

In the modern convex analysis the notion of strong convexity (concavity) is widely used [6]. Proposition 4 and inequality (1) imply the following observation.

COROLLARY 2. For arbitrary channel  $\Phi$  the function  $\chi_\Phi$  is a strong concave function on  $\mathfrak{S}(\mathcal{H})$  in the following sense:

$$\chi_\Phi(\lambda\rho_1 + (1-\lambda)\rho_2) \geq \lambda\chi_\Phi(\rho_1) + (1-\lambda)\chi_\Phi(\rho_2) + \frac{1}{2}\lambda(1-\lambda)\|\Phi(\bar{\rho}_2) - \Phi(\bar{\rho}_1)\|_1^2$$

for any states  $\rho_1$  and  $\rho_2$  in  $\mathfrak{S}(\mathcal{H})$  and any  $\lambda$  in  $[0, 1]$ .

Similarity of the properties of the functions  $\chi_\Phi(\rho)$  and  $H_\Phi(\rho)$  is stressed by the following analog of Simon's dominated convergence theorem for quantum entropy [33].<sup>4</sup> ■

COROLLARY 3. Let  $\{\rho_n\}$  be a sequence of states in  $\mathfrak{S}(\mathcal{H})$  converging to the state  $\rho_0$  such that  $\lambda_n \rho_n \leq \rho_0$  for some sequence  $\{\lambda_n\}$  of positive numbers converging to 1. Then

$$\lim_{n \rightarrow +\infty} \chi_\Phi(\rho_n) = \chi_\Phi(\rho_0).$$

*Proof.* The condition  $\lambda_n \rho_n \leq \rho_0$  implies decomposition  $\rho_0 = \lambda_n \rho_n + (1 - \lambda_n) \sigma_n$ , where  $\sigma_n = (1 - \lambda_n)^{-1}(\rho_0 - \lambda_n \rho_n)$  is a state in  $\mathfrak{S}(\mathcal{H})$ . By concavity of the  $\chi$ -function we have

$$\chi_\Phi(\rho_0) \geq \lambda_n \chi_\Phi(\rho_n) + (1 - \lambda_n) \chi_\Phi(\sigma_n) \geq \lambda_n \chi_\Phi(\rho_n) \quad \forall n,$$

<sup>4</sup>This theorem can be formulated as corollary 3 with the quantum entropy  $H$  instead of the function  $\chi_\Phi$ .

and hence  $\limsup_{n \rightarrow +\infty} \chi_{\Phi}(\rho_n) \leq \chi_{\Phi}(\rho_0)$ . Lower semicontinuity of the  $\chi$ -function implies existence of the above limit. Corollary 3 is proved.

*Remark 2.* Corollary 3 provides possibility to approximate the value  $\chi_{\Phi}(\rho_0)$  for arbitrary state  $\rho_0$  by the sequence  $\{\chi_{\Phi}(\rho_n)\}$ , in which  $\rho_n = (\text{Tr } P_n \rho_0)^{-1} P_n \rho_0$  is a finite rank state for each  $n$ , where  $P_n$  is the spectral projector of the state  $\rho_0$ , corresponding to  $n$  maximal eigenvalues.

By exploring the properties of the convex closure of the output entropy in the next section we will establish continuity of the restriction of the  $\chi$ -function to any set of continuity of the output entropy (theorem 1).

We shall also use the following chain properties of the  $\chi$ -function.

**PROPOSITION 5.** *Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  and  $\Psi: \mathfrak{S}(\mathcal{H}') \mapsto \mathfrak{S}(\mathcal{H}'')$  be two arbitrary channels. Then*

$$\chi_{\Psi \circ \Phi}(\rho) \leq \chi_{\Phi}(\rho) \quad \text{and} \quad \chi_{\Psi \circ \Phi}(\rho) \leq \chi_{\Psi}(\Phi(\rho)) \quad \text{for all } \rho \text{ in } \mathfrak{S}(\mathcal{H}).$$

*Proof.* The first inequality follows from the monotonicity property of the relative entropy [23] and definition (14), the second one is a direct corollary of definition (14).

**5. Convex closure of the output entropy.** In the finite dimensional case the output entropy  $H_{\Phi}$  is finite and the function  $\chi_{\Phi}$  can be represented by expression (15) as a difference between the output entropy  $H_{\Phi}$  and its convex hull  $\text{co } H_{\Phi}$ . In this case the function  $\text{co } H_{\Phi}$  is continuous and hence it is closed (see appendix A). This means that the convex hull  $\text{co } H_{\Phi}$  of the output entropy coincides with the convex closure  $\overline{\text{co}} H_{\Phi}$  of the output entropy.

In the infinite dimensional case the function  $\text{co } H_{\Phi}$  is not closed even in the case of the noiseless channel  $\Phi$ . Indeed,  $\text{co } H_{\Phi}(\rho) = +\infty$  for any state  $\rho$  such that  $H_{\Phi}(\rho) = +\infty$  (see the proof of lemma 2), but this state  $\rho$  is a limit of some sequence  $\{\rho_n\}$  of finite rank states, for which  $\text{co } H_{\Phi}(\rho_n) = 0$ . Thus the function  $\text{co } H_{\Phi}$  is not lower semicontinuous.

It seems natural to suppose that in the infinite dimensional case the role of the function  $\text{co } H_{\Phi}$  is played by the function  $\overline{\text{co}} H_{\Phi}$ . The aim of this section is to confirm this conjecture by exploring properties of the function  $\overline{\text{co}} H_{\Phi}(\rho)$  and its relation to the  $\chi$ -function.

First of all we will obtain the explicit representation for  $\overline{\text{co}} H_{\Phi}$ . Consider the function

$$\widehat{H}_{\Phi}(\rho) = \inf_{\pi \in \mathcal{P}_{\{\rho\}}} \int_{\mathfrak{S}(\mathcal{H})} H_{\Phi}(\rho) \pi(d\rho) \leq +\infty,$$

where  $\mathcal{P}_{\{\rho\}}$  is the set of all probability measures with the barycenter  $\rho$ . It is clear that  $\widehat{H}_{\Phi}(\rho) \leq \text{co } H_{\Phi}(\rho) \leq H_{\Phi}(\rho)$  for all states  $\rho$  in  $\mathfrak{S}(\mathcal{H})$ . By considering properties of the function  $\widehat{H}_{\Phi}$  we will establish that  $\widehat{H}_{\Phi} = \overline{\text{co}} H_{\Phi}$  (proposition 7).

In the previous section it is mentioned that in the definition of the  $\chi$ -function the supremum over all measures coincides with the supremum over all measures with finite support (conventional ensembles). In contrast to this in the case of the  $\widehat{H}$ -function we have the following observation.

**LEMMA 2.** *Equality  $\widehat{H}_{\Phi}(\rho) = \inf_{\sum_i \pi_i \rho_i = \rho} \sum_i \pi_i H_{\Phi}(\rho_i) = \text{co } H_{\Phi}(\rho)$  holds if and only if either  $H_{\Phi}(\rho) < +\infty$  or  $\widehat{H}_{\Phi}(\rho) = +\infty$ .*

*Proof.* If  $H_{\Phi}(\rho) < +\infty$  then  $\chi_{\Phi}(\rho) = H_{\Phi}(\rho) - \text{co } H_{\Phi}(\rho)$ . By [10, proposition 1 and corollary 1] we have  $\chi_{\Phi}(\rho) = H_{\Phi}(\rho) - \widehat{H}_{\Phi}(\rho)$  and hence  $\widehat{H}_{\Phi}(\rho) = \text{co } H_{\Phi}(\rho)$ .

If  $H_\Phi(\rho) = +\infty$  then  $\text{co} H_\Phi(\rho) = +\infty$  since the set of states with finite output entropy is convex [34]. Lemma 2 is proved.

Lemma 2 implies that  $\widehat{H}_\Phi(\rho) < \text{co} H_\Phi(\rho)$  for any state  $\rho$  such that  $H_\Phi(\rho) = +\infty$  and  $\widehat{H}_\Phi(\rho) < +\infty$ . Note that the set of such states is nonempty. For example, in the case of the noiseless channel  $\Phi$  it is easy to see that  $\widehat{H}_\Phi(\rho) = 0$  for any state  $\rho$ , but the set of all states  $\rho$  such that  $H_\Phi(\rho) < +\infty$  is a subset of the first category in the set of all states  $\mathfrak{S}(\mathcal{H})$  [34].

Show first that the infimum in the definition of the function  $\widehat{H}_\Phi(\rho)$  can be taken only over measures supported by pure states. Consider the following partial order on the set  $\mathcal{P}$ . Let  $\mathcal{S}$  be the set of all convex continuous bounded functions on the set  $\mathfrak{S}(\mathcal{H})$ . We say that  $\mu \succ \nu$  if and only if

$$\int_{\mathfrak{S}(\mathcal{H})} f(\rho) \mu(d\rho) \geq \int_{\mathfrak{S}(\mathcal{H})} f(\rho) \nu(d\rho) \quad \text{for all } f \text{ in } \mathcal{S}.$$

This partial order on the sets of probability measures on convex sets, often called the Choquet ordering, is studied in details (see., for example, [14]).

PROPOSITION 6. *For any state  $\rho_0$  there exists a measure  $\pi_0$  supported by the set of pure states with the barycenter  $\rho_0$  such that*

$$\widehat{H}_\Phi(\rho_0) = \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi_0(d\rho).$$

The measure  $\pi_0$  can be chosen to be a measure with support consisting of  $n^2$  atoms (ensemble of  $n^2$  pure states) if and only if the state  $\rho_0$  has finite rank  $n$ .

*Proof.* In the proof of the theorem in [10] it is shown that the functional

$$(18) \quad \pi \mapsto \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi(d\rho)$$

is well defined and lower semicontinuous on the set  $\mathcal{P}$  (endowed with the topology of weak convergence). By proposition 2 in [10] the set  $\mathcal{P}_{\{\rho_0\}}$  is compact in this topology. Hence this functional achieves its minimum on the set  $\mathcal{P}_{\{\rho_0\}}$  at some measures  $\pi_* \in \mathcal{P}_{\{\rho_0\}}$ , t.i.

$$(19) \quad \widehat{H}_\Phi(\rho_0) = \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi_*(d\rho).$$

To show that among all such measures  $\pi_*$  there exists a measure  $\pi_0$  supported by pure states we will use the following two simple properties of the above partial order.

1. Let  $\{\mu_n\}$  and  $\{\nu_n\}$  be two sequences in  $\mathcal{P}$ , weakly converging to measures  $\mu$  and  $\nu$  correspondingly, such that  $\mu_n \succ \nu_n$  for all  $n$ . Then  $\mu \succ \nu$ .
2. If  $\mu \succ \nu$  then

$$\int_{\mathfrak{S}(\mathcal{H})} g(\rho) \mu(d\rho) \geq \int_{\mathfrak{S}(\mathcal{H})} g(\rho) \nu(d\rho)$$

for any function  $g$ , which can be represented as a pointwise limit of monotonous sequence of functions in  $\mathcal{S}$ .

By lemma 1 in [10] there exists the sequence  $\{\pi_n\}$  of measures in  $\mathcal{P}_{\{\rho_0\}}$  with finite supports, weakly converging to  $\pi_*$ . Decomposing each atom of the measure  $\pi_n$  into

convex combination of pure states we obtain the measure  $\hat{\pi}_n$  with the same barycenter supported by the set of pure states. It is easy to see that  $\hat{\pi}_n \succ \pi_n$ . By compactness of the set  $\mathcal{P}_{\{\rho_0\}}$  there exists subsequence  $\{\hat{\pi}_{n_k}\}$  converging to some measure  $\pi_0$  supported by the set of pure states due to theorem 6.1 in [28]. Since  $\hat{\pi}_{n_k} \succ \pi_{n_k}$ , the above property 1 of the partial order  $\succ$  implies  $\pi_0 \succ \pi_*$ .

By lemma 4 in [22] the convex function  $g(\rho) = -H_\Phi(\rho) = -H(\Phi(\rho))$  is a pointwise limit of the monotonous sequence of bounded continuous functions

$$g_n(\rho) = -(\text{Tr } P_n \Phi(\rho)) H((\text{Tr } P_n \Phi(\rho))^{-1} P_n \Phi(\rho) P_n),$$

where  $\{P_n\}$  is an arbitrary sequence of finite dimensional projectors strongly increasing to the unit operator  $I$ . It is easy to see that the functions  $g_n$  are convex and hence lie in  $\mathcal{S}$  for all  $n$ . By the above property 2 of the partial order  $\succ$  (with  $g(\rho) = -H_\Phi(\rho)$ ) and (19) we have

$$\hat{H}_\Phi(\rho_0) = \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi_*(d\rho) \geq \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi_0(d\rho).$$

The definition of the function  $\hat{H}_\Phi$  implies equality in the above inequality.

Let us prove the last statement of the proposition. If the state  $\rho_0$  has infinite rank then the set  $\mathcal{P}_{\{\rho_0\}}$  contains no measures finitely supported by pure states.

Let the state  $\rho_0$  has finite rank  $n$ ,  $\mathcal{H}_0 = \text{supp } \rho_0$  be the  $n$ -dimensional subspace and  $\Phi_0$  be the subchannel of the channel  $\Phi$ , corresponding to the subspace  $\mathcal{H}_0$  (the subchannel  $\Phi_0$  of the channel  $\Phi$ , corresponding to the subspace  $\mathcal{H}_0$ , is the restriction of the channel  $\Phi$  to the set of states supported by the subspace  $\mathcal{H}_0$  [31]).

If  $H_{\Phi_0}(\rho_0) = H_\Phi(\rho_0) < +\infty$  then the function  $H_{\Phi_0}$  is continuous on the compact set  $\mathfrak{S}(\mathcal{H}_0)$  by the below lemma 3. Hence we can apply lemma A-2 in [35] to prove existence of ensemble consisting of  $(\dim \mathcal{H}_0)^2$  states with the average  $\rho_0$ , at which the infimum in the definition of the function  $\text{co } H_{\Phi_0}$ , coinciding with the restriction of the function  $\text{co } H_\Phi$  to the subset  $\mathfrak{S}(\mathcal{H}_0)$  of  $\mathfrak{S}(\mathcal{H})$ , is achieved. By lemma 2 the restriction of the function  $\text{co } H_\Phi$  to the set  $\mathfrak{S}(\mathcal{H}_0)$  coincides with the restriction of the function  $\hat{H}_\Phi$  to this set.

If  $H_{\Phi_0}(\rho_0) = H_\Phi(\rho_0) = +\infty$  then  $\hat{H}_\Phi(\rho_0) = +\infty$  and hence any ensemble with the average state  $\rho_0$  is optimal. To prove this note that  $H_\Phi(\rho_0) = +\infty$  implies  $H_\Phi(\sigma) = +\infty$  for any state  $\sigma$  such that  $\text{supp } \sigma = \text{supp } \rho_0 = \mathcal{H}_0$ . Indeed, for this state  $\sigma$  there exists a positive number  $\lambda_\sigma$  such that  $\lambda_\sigma \sigma \geq \rho_0$ . Nonnegativity of the relative entropy implies

$$\lambda_\sigma \text{Tr } \Phi(\sigma)(-\log \Phi(\sigma)) \geq \text{Tr } \Phi(\rho_0)(-\log \Phi(\sigma)) \geq \text{Tr } \Phi(\rho_0)(-\log \Phi(\rho_0)) = +\infty.$$

Suppose that  $\hat{H}_\Phi(\rho_0) < +\infty$ . Then there exists measure  $\pi$  with the barycenter  $\rho_0$  such that the function  $H_\Phi$  is finite  $\pi$ -almost everywhere. Let  $\mathcal{F}$  be subset of  $\mathfrak{S}(\mathcal{H}_0)$  such that the function  $H_\Phi$  is finite on the set  $\mathcal{F}$  and  $\pi(\mathcal{F}) = 1$ . The equality  $\rho_0 = \int_{\mathcal{F}} \rho \pi(d\rho)$  implies that the linear hull of subspaces  $\{\text{supp } \rho\}_{\rho \in \mathcal{F}}$  coincides with  $\mathcal{H}_0$  and hence there exists finite set  $\{\rho_i\}_{i=1}^n$  of states in  $\mathcal{F}$  such that  $\text{supp}(n^{-1} \sum_{i=1}^n \rho_i) = \mathcal{H}_0$ . Since the state  $n^{-1} \sum_{i=1}^n \rho_i$  is a finite convex combination of the states  $\rho_i$ ,  $i = 1, \dots, n$  such that  $H_\Phi(\rho_i) < +\infty$  for all  $i = 1, \dots, n$  we conclude that  $H_\Phi(n^{-1} \sum_{i=1}^n \rho_i) < +\infty$  [34]. But this contradicts to the previous observation. Proposition 6 is proved.

LEMMA 3. Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  be such channel that  $\dim \mathcal{H} < +\infty$ . If there exists a full rank state  $\rho_0$  such that  $H_\Phi(\rho_0) < +\infty$  then the function  $H_\Phi$  is continuous on the set  $\mathfrak{S}(\mathcal{H})$ .

*Proof.* Let  $I_{\mathcal{H}}$  be the unit operator in the space  $\mathcal{H}$ . Finite dimensionality of the space  $\mathcal{H}$  implies that  $\lambda I_{\mathcal{H}} \leq \rho_0$  for some positive  $\lambda$ , and hence  $H_{\Phi}(I_{\mathcal{H}}) < +\infty$ . The assertion of the lemma follows from Simon's dominated convergence theorem [33] (with using  $\Phi(I_{\mathcal{H}})$  in the role of the operator  $B$ ).

DEFINITION 4. A measure  $\pi_0$  with the properties stated in proposition 6 is called  $\widehat{H}_{\Phi}$ -optimal measure for the state  $\rho_0$ .

It is easy to see that the set of  $\widehat{H}_{\Phi}$ -optimal measures coincides with the set of  $\chi_{\Phi}$ -optimal measures for any state  $\rho_0$  with finite output entropy  $H_{\Phi}(\rho_0)$ .

The other important properties of the function  $\widehat{H}_{\Phi}(\rho)$  are established in the following lemma.

LEMMA 4. *The function  $\widehat{H}_{\Phi}(\rho)$  is convex and lower semicontinuous on the set  $\mathfrak{S}(\mathcal{H})$ .*

*Proof.* To prove convexity of the function  $\widehat{H}_{\Phi}$  it is sufficient to note that

$$\lambda \mathcal{P}_{\{\rho_1\}} + (1 - \lambda) \mathcal{P}_{\{\rho_2\}} \subseteq \mathcal{P}_{\{\lambda \rho_1 + (1 - \lambda) \rho_2\}}$$

for arbitrary states  $\rho_1, \rho_2$  and  $\lambda \in [0, 1]$ .

Suppose that the function  $\widehat{H}_{\Phi}$  is not lower semicontinuous. This implies existence of a sequence  $\{\rho_n\}$  converging to some state  $\rho_0$  such that

$$(20) \quad \lim_{n \rightarrow +\infty} \widehat{H}_{\Phi}(\rho_n) < \widehat{H}_{\Phi}(\rho_0).$$

By proposition 6 for each  $n = 1, 2, \dots$  there exists measure  $\pi_n$  in  $\mathcal{P}_{\{\rho_n\}}$  such that

$$\widehat{H}_{\Phi}(\rho_n) = \int_{\mathfrak{S}(\mathcal{H})} H_{\Phi}(\rho) \pi_n(d\rho).$$

The set  $\mathcal{A} = \{\rho_n\}_{n=0}^{+\infty}$  is a compact subset of  $\mathfrak{S}(\mathcal{H})$ . By proposition 2 in [10] the set  $\mathcal{P}_{\mathcal{A}}$  is compact. Since  $\{\pi_n\} \subset \mathcal{P}_{\mathcal{A}}$ , there exists subsequence  $\{\pi_{n_k}\}$  converging to some measure  $\pi_0$ . Continuity of the map  $\pi \mapsto \bar{\rho}(\pi)$  implies  $\pi_0 \in \mathcal{P}_{\{\rho_0\}}$ . By lower semicontinuity of the functional (18) we obtain

$$\begin{aligned} \widehat{H}_{\Phi}(\rho_0) &\leq \int_{\mathfrak{S}(\mathcal{H})} H_{\Phi}(\rho) \pi_0(d\rho) \leq \liminf_{k \rightarrow +\infty} \int_{\mathfrak{S}(\mathcal{H})} H_{\Phi}(\rho) \pi_{n_k}(d\rho) \\ &= \lim_{k \rightarrow +\infty} \widehat{H}_{\Phi}(\rho_{n_k}), \end{aligned}$$

which contradicts to (20). Lemma 4 is proved.

PROPOSITION 7. *The function  $\widehat{H}_{\Phi}$  coincides with the convex closure <sup>5</sup>  $\overline{\text{co}} H_{\Phi}$  of the output entropy  $H_{\Phi}$  and if  $\overline{\text{co}} H_{\Phi}(\rho) < +\infty$  then*

$$\{\overline{\text{co}} H_{\Phi}(\rho) = \text{co } H_{\Phi}(\rho)\} \iff \{H_{\Phi}(\rho) < +\infty\}.$$

*Proof.* Lemma 4 and the definition of the convex closure imply

$$(21) \quad \widehat{H}_{\Phi}(\rho) \leq \overline{\text{co}} H_{\Phi}(\rho) \leq \text{co } H_{\Phi}(\rho) \leq H_{\Phi}(\rho) \quad \forall \rho \in \mathfrak{S}(\mathcal{H}).$$

By lemma 2  $\widehat{H}_{\Phi}(\rho_0)$  coincides with  $\text{co } H_{\Phi}(\rho_0)$  for any state  $\rho_0$  with finite output entropy  $H_{\Phi}(\rho_0)$ . Thus (21) means that  $\widehat{H}_{\Phi}(\rho_0) = \overline{\text{co}} H_{\Phi}(\rho_0)$  for all such states.

<sup>5</sup>see. appendix A.

Let  $\rho_0$  be an arbitrary state such that  $\widehat{H}_\Phi(\rho_0) < +\infty$ . By the below lemma 5 there exists sequence  $\{\rho_n\}$  of states with finite output entropy, converging to the state  $\rho_0$ , such that  $\lim_{n \rightarrow +\infty} \widehat{H}_\Phi(\rho_n) = \widehat{H}_\Phi(\rho_0)$ . By the above observation  $\widehat{H}_\Phi(\rho_n) = \overline{\text{co}} H_\Phi(\rho_n)$  for all  $n$ . Since the function  $\overline{\text{co}} H_\Phi$  is lower semicontinuous (by definition), we obtain

$$\overline{\text{co}} H_\Phi(\rho_0) \leq \liminf_{n \rightarrow +\infty} \overline{\text{co}} H_\Phi(\rho_n) = \lim_{n \rightarrow +\infty} \widehat{H}_\Phi(\rho_n) = \widehat{H}_\Phi(\rho_0).$$

This inequality and (21) imply that  $\widehat{H}_\Phi(\rho_0) = \overline{\text{co}} H_\Phi(\rho_0)$  for arbitrary state  $\rho_0$ . Proposition 7 is proved.

The following observation plays essential role in study of the properties of the function  $\widehat{H}_\Phi$ .

LEMMA 5. *For arbitrary state  $\rho_0$  such that  $\widehat{H}_\Phi(\rho_0) < \infty$  there exists sequence  $\{\rho_n\}$  of finite rank states, converging to the state  $\rho_0$ , such that*

$$H_\Phi(\rho_n) < +\infty \quad \text{for all } n \quad \text{and} \quad \lim_{n \rightarrow +\infty} \widehat{H}_\Phi(\rho_n) = \widehat{H}_\Phi(\rho_0).$$

*Proof.* Let  $\pi_0$  be a  $\widehat{H}_\Phi$ -optimal measure for the state  $\rho_0$ , which exists by proposition 6. Since any probability measure on complete separable metric space  $\mathfrak{S}(\mathcal{H})$  is tight [1], [28], for any  $n \in \mathbf{N}$  there exists compact subset  $\mathcal{K}_n$  of the set  $\text{Extr}(\mathfrak{S}(\mathcal{H}))$  such that  $\pi_0(\mathcal{K}_n) > 1 - 1/n$ . Compactness of the set  $\mathcal{K}_n$  implies decomposition  $\mathcal{K}_n = \bigcup_{i=1}^{m(n)} \mathcal{A}_i^n$ , where  $\{\mathcal{A}_i^n\}_{i=1}^{m(n)}$  is a finite collection of disjoint measurable subsets with diameter less than  $1/n$ . Without loss of generality we may assume that  $\pi_0(\mathcal{A}_i^n) > 0$  for all  $i$  and  $n$ . By construction compact set  $\overline{\mathcal{A}_i^n}$  lies within some closed ball  $\mathcal{B}_i^n$  of diameter  $1/n$  for all  $i$  and  $n$ .

By assumption  $\widehat{H}_\Phi(\rho_0) = \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi_0(d\rho) < +\infty$ , and hence the function  $H_\Phi$  is finite  $\pi_0$ -almost everywhere. Since the function  $H_\Phi$  is lower semicontinuous it achieves its finite minimum on the compact set  $\overline{\mathcal{A}_i^n}$  of positive measure at some pure state  $\rho_i^n \in \overline{\mathcal{A}_i^n}$ . Consider the state  $\rho_n = (\pi_0(\mathcal{K}_n))^{-1} \sum_{i=1}^{m(n)} \pi_0(\mathcal{A}_i^n) \rho_i^n$ . We want to show that

$$(22) \quad \|\rho_n - \rho_0\|_1 \leq \frac{3}{n}.$$

The state  $\widehat{\rho}_i^n = (\pi_0(\mathcal{A}_i^n))^{-1} \int_{\mathcal{A}_i^n} \rho \pi_0(d\rho)$  lies in the set  $\mathcal{B}_i^n$  by convexity of  $\mathcal{B}_i^n$ . Hence  $\|\rho_i^n - \widehat{\rho}_i^n\|_1 \leq 1/n$ . By noting that  $\pi_0(\mathcal{K}_n) = \sum_{i=1}^{m(n)} \pi_0(\mathcal{A}_i^n)$ , we have

$$\begin{aligned} \|\rho_n - \rho_0\|_1 &= \left\| (\pi_0(\mathcal{K}_n))^{-1} \sum_{i=1}^{m(n)} \pi_0(\mathcal{A}_i^n) \rho_i^n \right. \\ &\quad \left. - \sum_{i=1}^{m(n)} \int_{\mathcal{A}_i^n} \rho \pi_0(d\rho) - \int_{\mathfrak{S}(\mathcal{H}) \setminus \mathcal{K}_n} \rho \pi_0(d\rho) \right\|_1 \\ &\leq \sum_{i=1}^{m(n)} \pi_0(\mathcal{A}_i^n) \|(\pi_0(\mathcal{K}_n))^{-1} \rho_i^n - \widehat{\rho}_i^n\|_1 + \left\| \int_{\mathfrak{S}(\mathcal{H}) \setminus \mathcal{K}_n} \rho \pi_0(d\rho) \right\|_1 \\ &\leq (1 - \pi_0(\mathcal{K}_n)) + \sum_{i=1}^{m(n)} \pi_0(\mathcal{A}_i^n) \|\rho_i^n - \widehat{\rho}_i^n\|_1 + \pi_0(\mathfrak{S}(\mathcal{H}) \setminus \mathcal{K}_n) < \frac{3}{n}, \end{aligned}$$

t.i. (22).



By the choice of the states  $\rho_i^n$  for each  $i$  and  $n$  we have  $H_\Phi(\rho_i^n) \leq H_\Phi(\rho)$  for all  $\rho$  in  $\mathcal{A}_i^n$ . Hence,

$$\begin{aligned} \widehat{H}_\Phi(\rho_n) &\leq (\pi_0(\mathcal{K}_n))^{-1} \sum_{i=1}^{m(n)} \pi_0(\mathcal{A}_i^n) H_\Phi(\rho_i^n) \\ &\leq (\pi_0(\mathcal{K}_n))^{-1} \sum_{i=1}^{m(n)} \int_{\mathcal{A}_i^n} H_\Phi(\rho) \pi_0(d\rho) \\ &\leq (\pi_0(\mathcal{K}_n))^{-1} \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi_0(d\rho) = (\pi_0(\mathcal{K}_n))^{-1} \widehat{H}_\Phi(\rho_0). \end{aligned}$$

Thus  $\limsup_{n \rightarrow +\infty} \widehat{H}_\Phi(\rho_n) \leq \widehat{H}_\Phi(\rho_0)$ . But  $\lim_{n \rightarrow +\infty} \rho_n = \rho_0$  due to (22) and by applying lemma 4 we obtain  $\liminf_{n \rightarrow +\infty} \widehat{H}_\Phi(\rho_n) \geq \widehat{H}_\Phi(\rho_0)$ . Hence there exists  $\lim_{n \rightarrow +\infty} \widehat{H}_\Phi(\rho_n) = \widehat{H}_\Phi(\rho_0)$ .

By construction the state  $\rho_n$  for each  $n$  is a finite convex combination of pure states  $\rho_i^n$ ,  $i = 1, \dots, m(n)$ , with finite output entropy  $H_\Phi(\rho_i^n)$ . It follows that  $H_\Phi(\rho_n) < +\infty$  for all  $n$  [34]. Lemma 5 is proved.

The real Banach space  $\mathfrak{B}_h(\mathcal{H})$  of all hermitian operators is a dual space for the space  $\mathfrak{T}_h(\mathcal{H})$  of all hermitian trace class operators. The nonnegative lower semicontinuous function  $H_\Phi$  on  $\mathfrak{S}(\mathcal{H})$  can be extended to the nonnegative lower semicontinuous function  $\overline{H}_\Phi$  on  $\mathfrak{T}_h(\mathcal{H})$  by ascribing the value  $+\infty$  to arbitrary operator in  $\mathfrak{T}_h(\mathcal{H}) \setminus \mathfrak{S}(\mathcal{H})$ . The Fenchel transform of the function  $\overline{H}_\Phi$  (see appendix A) is defined on the set  $\mathfrak{B}_h(\mathcal{H})$  by the expression

$$(23) \quad H_\Phi^*(A) = \sup_{\rho \in \mathfrak{T}_h(\mathcal{H})} (\text{Tr } A\rho - \overline{H}_\Phi(\rho)) = \sup_{\rho \in \mathfrak{S}(\mathcal{H})} (\text{Tr } A\rho - H_\Phi(\rho)).$$

The double Fenchel transform  $H_\Phi^{**}$  is defined on the set  $\mathfrak{T}_h(\mathcal{H})$  by the expression

$$(24) \quad H_\Phi^{**}(\rho) = \sup_{A \in \mathfrak{B}_h(\mathcal{H})} (\text{Tr } A\rho - H_\Phi^*(A)).$$

Since the function  $\overline{H}_\Phi$  is nonnegative, its convex closure  $\overline{\text{co}} \overline{H}_\Phi$  coincides with its double Fenchel transform  $H_\Phi^{**}$ . Since the restriction of the function  $\overline{\text{co}} \overline{H}_\Phi$  to the set  $\mathfrak{S}(\mathcal{H})$  coincides with  $\overline{\text{co}} H_\Phi$ , proposition 7 implies the following representation for the  $\widehat{H}$ -function.

COROLLARY 4. *Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  be a quantum channel. Then*

$$\widehat{H}_\Phi(\rho) = H_\Phi^{**}(\rho) = \sup_{A \in \mathfrak{B}_h(\mathcal{H})} \inf_{\sigma \in \mathfrak{S}(\mathcal{H})} (H_\Phi(\sigma) + \text{Tr } A(\rho - \sigma))$$

for any state  $\rho \in \mathfrak{S}(\mathcal{H})$ .

Consider the set  $\widehat{H}_\Phi^{-1}(0) = \{\rho \in \mathfrak{S}(\mathcal{H}) \mid \widehat{H}_\Phi(\rho) = 0\}$ . Note that the set  $H_\Phi^{-1}(0) = \{\rho \in \mathfrak{S}(\mathcal{H}) \mid H_\Phi(\rho) = 0\}$  is a closed subset of  $\mathfrak{S}(\mathcal{H})$  due to lower semicontinuity of the quantum entropy [34].

PROPOSITION 8. *The set  $\widehat{H}_\Phi^{-1}(0)$  coincides with the convex closure of the set  $H_\Phi^{-1}(0) \cap \text{Extr } \mathfrak{S}(\mathcal{H})$ .*

*Proof.* Let  $\rho_0 \in \overline{\text{co}}(H_\Phi^{-1}(0) \cap \text{Extr } \mathfrak{S}(\mathcal{H}))$ . Then there exists sequence of states  $\{\rho_n\} \subset \text{co}(H_\Phi^{-1}(0) \cap \text{Extr } \mathfrak{S}(\mathcal{H}))$ , converging to the state  $\rho_0$ . By definition  $\widehat{H}_\Phi(\rho_n) = 0$ . Nonnegativity and lower semicontinuity of the function  $\widehat{H}_\Phi$  (lemma 4) implies  $\widehat{H}_\Phi(\rho_0) = 0$ .

Let  $\rho_0 \in \widehat{H}_\Phi^{-1}(0)$ . By proposition 6 the state  $\rho_0$  is the barycenter of some measure  $\pi_0$  supported by pure states such that  $H_\Phi(\rho) = 0$  for  $\pi_0$ -almost all  $\rho$ . By using the arguments from the proof of theorem 6.3 in [28] it is easy to see that this measure  $\pi_0$  can be approximated by the sequence of measures  $\pi_n$ , finitely supported by the set of pure states, such that  $H_\Phi(\rho) = 0$  for  $\pi_n$ -almost all  $\rho$ . It follows that for each  $n$  all atoms of the measure  $\pi_n$  are pure states in  $H_\Phi^{-1}(0)$ . By continuity of the map  $\pi \mapsto \bar{\rho}(\pi)$  the state  $\rho_0 = \bar{\rho}(\pi_0)$  is a limit of the sequence  $\{\bar{\rho}(\pi_n)\}$  of states in  $\text{co}(H_\Phi^{-1}(0) \cap \text{Extr } \mathfrak{S}(\mathcal{H}))$ . Proposition 8 is proved.

**6. On continuity of the functions  $\chi_\Phi$  and  $\widehat{H}_\Phi$ .** Lemma 2 implies

$$(25) \quad \chi_\Phi(\rho) = H_\Phi(\rho) - \widehat{H}_\Phi(\rho)$$

for all states  $\rho$  with finite output entropy. This expression remains valid in the case  $H_\Phi(\rho) = +\infty$  and  $\widehat{H}_\Phi(\rho) < +\infty$ . Indeed, by substituting  $\widehat{H}_\Phi$ -optimal measure  $\pi$  for the state  $\rho$  into expression (4) in [10] it is easy to obtain that  $\chi_\Phi(\rho) = +\infty$ . Note also that for any state  $\rho$  finiteness of  $\chi_\Phi(\rho)$  and  $\widehat{H}_\Phi(\rho)$  implies finiteness of  $H_\Phi(\rho)$  and hence validity of expression (25). The last assertion can be proved by using proposition 4 and lemma 5.

By expression (25) continuity of the functions  $\chi_\Phi$  and  $\widehat{H}_\Phi$  on some set  $\mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$  implies continuity of the output entropy  $H_\Phi$  on this set. It is essential that the converse assertion also holds and follows from expression (25) due to lower semicontinuity of the function  $\chi_\Phi$  (proposition 4) and lower semicontinuity of the function  $\widehat{H}_\Phi$  (lemma 4).

**THEOREM 1.** *Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  be a quantum channel. If the restriction of the output entropy  $H_\Phi$  to some set  $\mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$  is continuous then the restrictions of the functions  $\chi_\Phi$  and  $\widehat{H}_\Phi$  to the set  $\mathcal{A}$  are also continuous.*

*Remark 3.* The statement of theorem 1 seems surprising by the following reason. The value of the output entropy  $H_\Phi$  at a particular state  $\rho$  is completely defined by the output state  $\Phi(\rho)$  and it does not depend on the action of the channel  $\Phi$  on other input states. Thus continuity of the function  $H_\Phi$  on some set  $\mathcal{A}$  means continuity of the entropy on the set  $\Phi(\mathcal{A})$ , which depends only on the set  $\Phi(\mathcal{A})$ . In contrast to this the values of the functions  $\chi_\Phi$  and  $\widehat{H}_\Phi$  at some state  $\rho_0$  depend (due to their definitions) on the action of the channel  $\Phi$  to all states contained in the union of supports of all measures with the barycenter  $\rho_0$ . Thus behavior of these functions on some set  $\mathcal{A}$  depends on action of the channel  $\Phi$  to all states contained in the union of supports of all measures with the barycenter in  $\mathcal{A}$ . Nevertheless, by theorem 1 continuity of the entropy on the set  $\Phi(\mathcal{A})$  guarantees continuity of the functions  $\chi_\Phi$  and  $\widehat{H}_\Phi$  on the set  $\mathcal{A}$  independently of the action of the channel  $\Phi$  to all states, which are not contained in  $\mathcal{A}$ .

Note also that continuity of one of the functions  $\chi_\Phi$  and  $\widehat{H}_\Phi$  on some set  $\mathcal{A}$  does not imply continuity of the output entropy on this set. For example, in the case of the noiseless channel  $\Phi$  the function  $\widehat{H}_\Phi$  is equal to zero on  $\mathfrak{S}(\mathcal{H})$ , but the function  $H_\Phi$  (the entropy of a state) is discontinuous on  $\mathfrak{S}(\mathcal{H})$ .

Proposition 3 in [10] implies the following observation.

**COROLLARY 5.** *Let  $H'$  be positive unbounded operator in the space  $\mathcal{H}'$  such that  $\text{Tr exp}(-\beta H') < +\infty$  for all  $\beta > 0$ .<sup>6</sup> Then the restrictions of the functions  $\chi_\Phi$  and*

<sup>6</sup>In [10] such operator is called  $\mathfrak{H}$ -operator.

$\widehat{H}_\Phi$  to the subset

$$\mathcal{A}_{h'} = \{\rho \in \mathfrak{S}(\mathcal{H}) \mid \text{Tr } \Phi(\rho) H' \leq h'\}$$

are continuous for each  $h' \geq 0$ .

In [10] it is mentioned that the condition of corollary 5 is fulfilled for quantum Gaussian channels with the energy constraint  $\text{Tr } \rho H \leq h$ , where  $H = R^\top \varepsilon R$  is the Hamiltonian of a system of oscillators with nongenerate energy matrix  $\varepsilon$  and  $R$  are the canonical variables of the system.

**PROPOSITION 9.** *Let  $\{\rho_n\}$  be a sequence of states converging to the state  $\rho_0$  such that  $\lim_{n \rightarrow +\infty} \widehat{H}_\Phi(\rho_n) = \widehat{H}_\Phi(\rho_0)$ <sup>7</sup> and  $\pi_n$  be a  $\widehat{H}_\Phi$ -optimal measure for the state  $\rho_n$  for each  $n = 1, 2, \dots$ . The set of partial limits of the sequence  $\{\pi_n\}_{n=1}^{+\infty}$  is nonempty and consists of  $\widehat{H}_\Phi$ -optimal measures for the state  $\rho_0$ .*

*Proof.* Let  $\{\rho_n\}$  and  $\{\pi_n\}$  be the above sequences. Since the set  $\{\rho_n\}_{n=0}^{+\infty}$  is a compact subset of  $\mathfrak{S}(\mathcal{H})$ , the set  $\mathcal{P}_{\{\{\rho_n\}_{n=0}^{+\infty}\}}$  is a compact subset of  $\mathcal{P}$  by proposition 2 in [10]. Hence the sequence  $\{\pi_n\} \subseteq \mathcal{P}_{\{\{\rho_n\}_{n=0}^{+\infty}\}}$  has partial limits. Let  $\pi_0$  be a limit of some subsequence  $\{\pi_{n_k}\}$  of the sequence  $\{\pi_n\}$ . By lower semicontinuity of functional (18) we have

$$\begin{aligned} \widehat{H}_\Phi(\rho_0) &= \lim_{k \rightarrow +\infty} \widehat{H}_\Phi(\rho_{n_k}) = \lim_{k \rightarrow +\infty} \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi_{n_k}(d\rho) \\ &\geq \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi_0(d\rho), \end{aligned}$$

which means that  $\pi_0$  is a  $\widehat{H}_\Phi$ -optimal measure for the state  $\rho_0$ . Proposition 9 is proved.

Theorem 1 and proposition 9 make possible to represent  $\widehat{H}_\Phi$ -optimal ( $\chi_\Phi$ -optimal) measure for any state with finite output entropy as a limit point of the sequence of measures with finite support (conventional ensembles). Recall that by proposition 6 for any state of finite rank  $n$  there exists  $\widehat{H}_\Phi$ -optimal measure with finite support consisting of  $n^2$  atoms (pure states).

**COROLLARY 6.** *Let  $\rho_0$  be such state that  $H_\Phi(\rho_0) < +\infty$ ,  $P_n$  be the spectral projector of  $\rho_0$  corresponding to its  $n$  maximal eigenvalues and  $\pi_n$  be a  $\widehat{H}_\Phi$ -optimal ( $\chi_\Phi$ -optimal) measure with finite support (ensemble of  $n^2$  pure states) for the finite rank state  $\rho_n = (\text{Tr } P_n \rho_0)^{-1} P_n \rho_0$  for each  $n \in \mathbf{N}$ . Then any partial limit of the sequence  $\{\pi_n\}$  is a  $\widehat{H}_\Phi$ -optimal measure for the state  $\rho_0$ .*

*Proof.* By using Simon's dominated convergence theorem [33] it is easy to show that

$$\lim_{n \rightarrow +\infty} H_\Phi(\rho_n) = H_\Phi(\rho_0).$$

Hence the conditions of theorem 1 are valid for the set  $\mathcal{A} = \{\rho_n\}_{n=0}^{+\infty}$ . Thus  $\lim_{n \rightarrow +\infty} \widehat{H}_\Phi(\rho_n) = \widehat{H}_\Phi(\rho_0)$ , and proposition 9 implies the assertion of the corollary.

<sup>7</sup>For example, the sequence of finite rank states provided by lemma 5.

**7. The additivity problem.** Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  and  $\Psi: \mathfrak{S}(\mathcal{K}) \mapsto \mathfrak{S}(\mathcal{K}')$  be channels with the constraints defined by the sets  $\mathcal{A}$  and  $\mathcal{B}$  correspondingly. Consider the channel  $\Phi \otimes \Psi: \mathfrak{S}(\mathcal{H} \otimes \mathcal{K}) \mapsto \mathfrak{S}(\mathcal{H}' \otimes \mathcal{K}')$ . For this channel it is natural to consider constraint defined by the set

$$\mathcal{A} \otimes \mathcal{B} = \{\omega \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K}) \mid \omega^{\mathcal{H}} \in \mathcal{A}, \omega^{\mathcal{K}} \in \mathcal{B}\},$$

where the following notations  $\omega^{\mathcal{H}} = \text{Tr}_{\mathcal{K}} \omega$  and  $\omega^{\mathcal{K}} = \text{Tr}_{\mathcal{H}} \omega$  are used.

Additivity of the Holevo capacity for the  $\mathcal{A}$ -constrained channel  $\Phi$  and the  $\mathcal{B}$ -constrained channel  $\Psi$  means equality

$$(26) \quad \overline{C}(\Phi \otimes \Psi, \mathcal{A} \otimes \mathcal{B}) = \overline{C}(\Phi, \mathcal{A}) + \overline{C}(\Psi, \mathcal{B}).$$

*Remark 4.* Let  $\Omega(\Phi, \mathcal{A})$  and  $\Omega(\Psi, \mathcal{B})$  be the output optimal average states for the  $\mathcal{A}$ -constrained channel  $\Phi$  and the  $\mathcal{B}$ -constrained channel  $\Psi$  correspondingly. Additivity of the Holevo capacity (26) implies that the state  $\Omega(\Phi, \mathcal{A}) \otimes \Omega(\Psi, \mathcal{B})$  is the output optimal average for the  $\mathcal{A} \otimes \mathcal{B}$ -constrained channel  $\Phi \otimes \Psi$ . Indeed, let  $\{\{\pi_i^k, \rho_i^k\}\}_k$  and  $\{\{\mu_j^k, \sigma_j^k\}\}_k$  be approximating sequences of ensembles for the  $\mathcal{A}$ -constrained channel  $\Phi$  and the  $\mathcal{B}$ -constrained channel  $\Psi$ . By proposition 1 the sequences  $\{\Phi(\bar{\rho}_k)\}_k$  and  $\{\Psi(\bar{\sigma}_k)\}_k$  converge to the states  $\Omega(\Phi, \mathcal{A})$  and  $\Omega(\Psi, \mathcal{B})$  correspondingly. It follows from (26) that the sequence of ensembles  $\{\{\pi_i^k \mu_j^k, \rho_i^k \otimes \sigma_j^k\}\}_k$  is approximating for the  $\mathcal{A} \otimes \mathcal{B}$ -constrained channel  $\Phi \otimes \Psi$ . By proposition 1 the limit  $\Omega(\Phi, \mathcal{A}) \otimes \Omega(\Psi, \mathcal{B})$  of the sequence  $\{\Phi(\bar{\rho}_k) \otimes \Psi(\bar{\sigma}_k)\}_k$  is the output optimal average state for the  $\mathcal{A} \otimes \mathcal{B}$ -constrained channel  $\Phi \otimes \Psi$ .

Additivity of the Holevo capacity (26) for arbitrary sets  $\mathcal{A}$  and  $\mathcal{B}$  is equivalent to validity of the following inequality:

$$(27) \quad \chi_{\Phi \otimes \Psi}(\omega) \leq \chi_{\Phi}(\omega^{\mathcal{H}}) + \chi_{\Psi}(\omega^{\mathcal{K}})$$

for any state  $\omega \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$  (subadditivity of the  $\chi$ -function). Indeed, let  $\omega$  be arbitrary state in  $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ . Let  $\mathcal{A} = \{\omega^{\mathcal{H}}\}$  and  $\mathcal{B} = \{\omega^{\mathcal{K}}\}$ . Then (26) and the definition of the  $\chi$ -function imply

$$\begin{aligned} \chi_{\Phi}(\omega^{\mathcal{H}}) + \chi_{\Psi}(\omega^{\mathcal{K}}) &= \overline{C}(\Phi, \{\omega^{\mathcal{H}}\}) + \overline{C}(\Psi, \{\omega^{\mathcal{K}}\}) \\ &= \overline{C}(\Phi \otimes \Psi, \{\omega^{\mathcal{H}}\} \otimes \{\omega^{\mathcal{K}}\}) \geq \chi_{\Phi \otimes \Psi}(\omega). \end{aligned}$$

Conversely, it follows from (27) that

$$\overline{C}(\Phi \otimes \Psi, \mathcal{A} \otimes \mathcal{B}) \leq \overline{C}(\Phi, \mathcal{A}) + \overline{C}(\Psi, \mathcal{B}).$$

Since the converse inequality is obvious we obtain (26). In [31] it is shown that subadditivity of the  $\chi$ -function holds for nontrivial class of infinite dimensional channels. ■

In [13] the approach to the additivity problem in the finite dimensional case based on the convex analysis is proposed. The results of the previous sections make possible to generalize this approach to the case of infinite dimensional channels.

For channel  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  and operator  $A \in \mathfrak{B}_+(\mathcal{H})$  consider the following characteristic [30]:

$$(28) \quad \nu_H(\Phi, A) = \inf_{\rho \in \mathfrak{S}(\mathcal{H})} (H_{\Phi}(\rho) + \text{Tr} A\rho).$$

Note that this characteristic is a generalization of the minimal output entropy of the channel  $\Phi$ , defined by the expression

$$(29) \quad H_{\min}(\Phi) = \inf_{\rho \in \mathfrak{S}(\mathcal{H})} H_{\Phi}(\rho) = \nu_H(\Phi, 0).$$

Concavity of the quantum entropy implies that the infimum in (28) and (29) can be taken over the set of all pure states  $\rho$  in  $\mathfrak{S}(\mathcal{H})$ .

Additivity of the minimal output entropy for the channels  $\Phi$  and  $\Psi$  means equality

$$(30) \quad H_{\min}(\Phi \otimes \Psi) = H_{\min}(\Phi) + H_{\min}(\Psi),$$

which is a partial case of additivity of the above characteristic with respect to the Kronecker sum:

$$(31) \quad \nu_H(\Phi \otimes \Psi, A \otimes I + I \otimes B) = \nu_H(\Phi, A) + \nu_H(\Psi, B).$$

If  $\Phi$  and  $\Psi$  are finite dimensional channels then in [21] it is shown that validity of inequality (27) for all states  $\omega$  is equivalent to validity of the following inequality

$$(32) \quad \widehat{H}_{\Phi \otimes \Psi}(\omega) \geq \widehat{H}_{\Phi}(\omega^{\mathcal{H}}) + \widehat{H}_{\Psi}(\omega^{\mathcal{K}})$$

for all states  $\omega \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$  (superadditivity of the  $\widehat{H}$ -function). In the infinite dimensional case the relations between these properties of sub-(super-)additivity are presented in the following theorem.

Let  $\mathfrak{S}_{\Phi, \Psi} = \{\omega \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K}) \mid H_{\Phi}(\omega^{\mathcal{H}}) < +\infty, H_{\Psi}(\omega^{\mathcal{K}}) < +\infty\}$  be a convex subset of  $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ .

**THEOREM 2.** *Let  $\Phi: \mathfrak{S}(\mathcal{H}) \mapsto \mathfrak{S}(\mathcal{H}')$  and  $\Psi: \mathfrak{S}(\mathcal{K}) \mapsto \mathfrak{S}(\mathcal{K}')$  be arbitrary channels.*

1) *Validity of inequality (32) for all states  $\omega \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$  is equivalent to validity of equality (31) for all positive operators  $A \in \mathfrak{B}_+(\mathcal{H})$  and  $B \in \mathfrak{B}_+(\mathcal{K})$ , which implies additivity of the minimal output entropy (30).*

2) *Validity of inequality (32) for all states  $\omega \in \mathfrak{S}_{\Phi, \Psi}$  is equivalent to validity of inequality (27) for all states  $\omega \in \mathfrak{S}_{\Phi, \Psi}$ , which means additivity of the Holevo capacity (26) for all sets  $\mathcal{A}$  and  $\mathcal{B}$  such that*

$$H_{\Phi}(\rho) < +\infty \quad \forall \rho \in \mathcal{A} \quad \text{and} \quad H_{\Psi}(\sigma) < +\infty \quad \forall \sigma \in \mathcal{B}.$$

*Proof.* 1) By considering product states it is easy to obtain the following subadditivity property:

$$(33) \quad \nu_H(\Phi \otimes \Psi, A \otimes I + I \otimes B) \leq \nu_H(\Phi, A) + \nu_H(\Psi, B).$$

By proposition 7 the function  $\widehat{H}_{\Phi}$  is the convex closure of the function  $H_{\Phi}$ . The Fenchel transform  $H_{\Phi}^*$  of the function  $H_{\Phi}$  is defined on the set  $\mathfrak{B}_h(\mathcal{H})$  of all hermitian operators by expression (23). By lemma 1 in [13] (formally the considered functions do not satisfy the conditions of this lemma, but it is easy to see that all arguments in its proof remain valid in this case), superadditivity of the  $\widehat{H}$ -function is equivalent to subadditivity of the Fenchel transform with respect to the Kronecker sum:

$$\begin{aligned} H_{\Phi \otimes \Psi}^*(A \otimes I_{\mathcal{K}} + I_{\mathcal{H}} \otimes B) &\leq H_{\Phi}^*(A) + H_{\Psi}^*(B) \\ \forall A \in \mathfrak{B}_h(\mathcal{H}), \quad \forall B \in \mathfrak{B}_h(\mathcal{K}). \end{aligned}$$

By definition the last inequality means the following one:

$$\begin{aligned} &\sup_{\omega \in \mathfrak{S}(\mathcal{H} \otimes \mathcal{K})} (\text{Tr } A\omega^{\mathcal{H}} + \text{Tr } B\omega^{\mathcal{K}} - H(\Phi \otimes \Psi(\omega))) \\ &\leq \sup_{\rho \in \mathfrak{S}(\mathcal{H})} (\text{Tr } A\rho - H(\Phi(\rho))) + \sup_{\sigma \in \mathfrak{S}(\mathcal{K})} (\text{Tr } B\sigma - H(\Psi(\sigma))) \end{aligned}$$

for all  $A \in \mathfrak{B}_h(\mathcal{H})$  and  $B \in \mathfrak{B}_h(\mathcal{K})$ .

By using invariance of the last inequality with respect to substitution for  $A$  and  $B$  by  $A \pm \|A\|I_{\mathcal{H}}$  and  $B \pm \|B\|I_{\mathcal{K}}$  correspondingly and using (33), we obtain assertion 1) of the theorem.

2) By representation (25) and subadditivity of the quantum entropy, inequality (32) for arbitrary state  $\omega \in \mathfrak{S}_{\Phi, \Psi}$  implies inequality (27) for this state.

Suppose that inequality (27) holds for any state  $\omega \in \mathfrak{S}_{\Phi, \Psi}$ . By the observation after remark 4 this means additivity of the Holevo capacity (26) for all  $\mathcal{A}$  and  $\mathcal{B}$  such that  $H_{\Phi}(\rho) < +\infty$  for all  $\rho \in \mathcal{A}$  and  $H_{\Psi}(\sigma) < +\infty$  for all  $\sigma \in \mathcal{B}$ . In particular,  $\overline{C}(\Phi \otimes \Psi, \{\omega^{\mathcal{H}}\} \otimes \{\omega^{\mathcal{K}}\}) = \overline{C}(\Phi, \{\omega^{\mathcal{H}}\}) + \overline{C}(\Psi, \{\omega^{\mathcal{K}}\})$ .

By remark 4 the state  $\Phi(\omega^{\mathcal{H}}) \otimes \Psi(\omega^{\mathcal{K}})$  is the output optimal average for the  $\{\omega^{\mathcal{H}}\} \otimes \{\omega^{\mathcal{K}}\}$ -constrained channel  $\Phi \otimes \Psi$ . By noting that  $\omega \in \{\omega^{\mathcal{H}}\} \otimes \{\omega^{\mathcal{K}}\}$  and using proposition 3 we obtain

$$\begin{aligned} \chi_{\Phi}(\omega^{\mathcal{H}}) + \chi_{\Psi}(\omega^{\mathcal{K}}) &= \overline{C}(\Phi, \{\omega^{\mathcal{H}}\}) + \overline{C}(\Psi, \{\omega^{\mathcal{K}}\}) \\ &= \overline{C}(\Phi \otimes \Psi, \{\omega^{\mathcal{H}}\} \otimes \{\omega^{\mathcal{K}}\}) \\ (34) \qquad \qquad \qquad &\geq \chi_{\Phi \otimes \Psi}(\omega) + H((\Phi \otimes \Psi)(\omega) \| \Phi(\omega^{\mathcal{H}}) \otimes \Psi(\omega^{\mathcal{K}})). \end{aligned}$$

Since

$$\begin{aligned} &H((\Phi \otimes \Psi)(\omega) \| \Phi(\omega^{\mathcal{H}}) \otimes \Psi(\omega^{\mathcal{K}})) \\ &= H(\Phi(\omega^{\mathcal{H}})) + H(\Psi(\omega^{\mathcal{K}})) - H((\Phi \otimes \Psi)(\omega)), \end{aligned}$$

inequality (34) coming with (25) leads to (32). Theorem 2 is proved.

In contrast to the finite dimensional case theorem 2 does not allow to show that subadditivity of the  $\chi$ -function (validity of inequality (27) for all states  $\omega$ ) implies superadditivity of the  $\widehat{H}$ -function (validity of inequality (32) for all states  $\omega$ ) for given two channels and vice versa. But for the nontrivial class of channels with finite output entropy (see the examples in [3]) theorem 2 guarantees equivalence of these properties.

**COROLLARY 7.** *For arbitrary two channels  $\Phi$  and  $\Psi$  with finite output entropy subadditivity of the  $\chi$ -function is equivalent to superadditivity of the  $\widehat{H}$ -function.*

The main difficulty preventing to prove the analogous assertion for arbitrary channels is related with existence of “superentangled” pure states  $\omega$ , having partial states with infinite entropy (see remark 4 in [31]).

**8. On definition of the Entanglement of Formation.** Entanglement is a specific feature of composed quantum systems. One of the measures of entanglement of a state of a bipartite system is the Entanglement of Formation (EoF) [17]. In the finite dimensional case it is defined as follows:

$$E_F(\rho) = \min_{\sum_i \pi_i \rho_i = \rho} \sum_i \pi_i H_{\Phi}(\rho_i),$$

where  $\Phi$  is the partial trace considered as a channel from the state space of composed system into the state space of its subsystem. In terms of convex analysis this definition means that the EoF coincides with the convex hull of the output entropy of the partial trace channel. Continuity of the EoF established in [26] implies that it coincides with the convex closure of the output entropy of the partial trace channel in this case.

The following generalization of the EoF to the infinite dimensional case is considered in [19]:

$$E_F^D(\rho) = \inf_{\sum_i \pi_i \rho_i = \rho} \sum_i \pi_i H_\Phi(\rho_i),$$

where the infimum is over all countable decompositions of the state  $\rho$  into pure states and  $\Phi$  is the partial trace channel.

The alternative approach to the definition of the EoF is considered in [24] in the case of tensor product of two systems with one of them finite dimensional. By using the results of the previous section we can generalize this approach and define the EoF in the general case by

$$E_F^C(\rho) = \widehat{H}_\Phi(\rho) = \inf_{\pi \in \mathcal{P}_{\{\rho\}}} \int_{\mathfrak{S}(\mathcal{H})} H_\Phi(\rho) \pi(d\rho),$$

where  $\Phi$  is the partial trace.

Proposition 7 shows that  $E_F^C$  is a convex lower semicontinuous function, coinciding with the convex closure of the output entropy of the partial trace channel. Proposition 6 implies that the infimum in the above expression is achieved at some measure supported by pure states. Proposition 8 guarantees the following natural property of  $E_F^C$ :

$$\{E_F^C(\rho) = 0\} \iff \{\text{state } \rho \text{ is separable}\},$$

where the set of separable (nonentangled) states is defined as the convex closure of pure product states [3]. Indeed, if  $\Phi$  is the partial trace channel then the set  $H_\Phi^{-1}(0) \cap \text{Extr } \mathfrak{S}(\mathcal{H})$  coincides with the set of pure product states. Theorem 1 guarantees continuity of  $E_F^C$  on subsets of states, on which the output entropy of one of the partial traces is continuous. By proposition 3 in [19] this implies continuity of  $E_F^C$  on the subsets of states with constrained mean energy. Note also that theorem 1 implies continuity of  $E_F^C = E_F^D$  on the whole state space of composed system, which contains at least one finite dimensional subsystem. (Proof of continuity of the EoF is nontrivial even in the finite dimensional case [26].)

The interesting question is the relations between  $E_F^D$  and  $E_F^C$ . Proposition 6 implies

$$E_F^D(\rho) \geq E_F^C(\rho)$$

for all states  $\rho$ . Since an arbitrary state can be represented as a countable convex combination of pure states it follows from lemma 2 and concavity of the output entropy that

$$(35) \quad E_F^D(\rho) = E_F^C(\rho)$$

for all states  $\rho$  having at least one partial trace with finite entropy. It is easy to see that (35) holds for all nonentangled and all pure states (for which  $\widehat{H}_\Phi$  coincides with  $H_\Phi$ ). Note that lemma 5 implies

$$E_F^C(\rho) = \lim_{\varepsilon \rightarrow +0} \inf_{\sum_i \pi_i \rho_i \in \mathcal{U}_\varepsilon(\rho)} \sum_i \pi_i H_\Phi(\rho_i),$$

where  $\mathcal{U}_\varepsilon(\rho)$  is  $\varepsilon$ -vicinity of the state  $\rho$  and the infimum is over all (finite) ensembles of pure states. But validity of equality (35) for mixed states having partial traces with infinite entropy remains an open problem. In the appendix B it is shown that validity of this equality can not be proved by using only general functional properties of the output entropy.

## 9. Appendixes.

**A. Convex hull and convex closure.** Here the notions from the convex analysis used in the main text are presented, following [4]. Let  $f$  be an arbitrary real valued function defined on closed convex subset  $X$  of some locally convex linear topological space. Consider the subset  $\text{epi}(f) = \{(x, \lambda) \in X \times \mathbf{R} \mid \lambda \geq f(x)\} \subseteq X \times \mathbf{R}$ .

Note that a function  $f$  is uniquely determined by the corresponding set  $\text{epi}(f)$ . Function  $f$  is called *convex* if the set  $\text{epi}(f)$  is a convex subset of  $X \times \mathbf{R}$ . Function  $f$  is called *closed* if the set  $\text{epi}(f)$  is a closed subset of  $X \times \mathbf{R}$ . Function  $f$  is called *proper* if it does not take the value  $-\infty$ . For proper function  $f$  convexity means

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad \forall x_1, x_2 \in X, \quad \forall \lambda \in [0, 1].$$

Each closed function  $f$  is lower semicontinuous in the sense that the set defined by the inequality  $f(x) \leq \lambda$  is a closed subset of  $X$  for arbitrary  $\lambda \in \mathbf{R}$  and, conversely, each lower semicontinuous function  $f$  is closed. It is possible to show that lower semicontinuity of a function  $f$  means that

$$\liminf_{n \rightarrow +\infty} f(x_n) \geq f(x_0)$$

for any sequence  $\{x_n\} \subset X$  converging to  $x_0$ .

Let  $f$  be an arbitrary function on  $X$ . The convex hull  $\text{co } f$  of the function  $f$  is defined by the expression

$$\text{co } f(x) = \inf_{(x, \lambda) \in \text{co}(\text{epi}(f))} \lambda,$$

in which symbol  $\text{co}$  in the right side denotes the convex hull of a set. This is equivalent to the following representation:

$$\text{co } f(x) = \inf_{\sum_i \pi_i x_i = x} \sum_i \pi_i f(x_i), \quad \pi_i > 0, \quad \sum_i \pi_i = 1.$$

It follows that  $\text{co } f$  is the greatest convex function majorized by  $f$ . The convex closure  $\overline{\text{co}} f$  of the function  $f$  is defined by the relation  $\text{epi}(\overline{\text{co}} f) = \overline{\text{co}}(\text{epi}(f))$ , in which symbol  $\overline{\text{co}}$  in the right side denotes the convex closure of a set. Hence  $\overline{\text{co}} f$  is the greatest convex closed function majorized by  $f$ . This implies:

$$\overline{\text{co}} f(x) \leq \text{co } f(x) \leq f(x) \quad \forall x \in X.$$

If  $f$  is a continuous function on *compact* convex set  $X$  then  $\overline{\text{co}} f = \text{co } f$  [12].

For arbitrary real valued function  $f$  on locally convex real linear topological space  $X$  the Fenchel transform  $f^*$  is the function on the dual space  $X^*$  defined by the expression

$$f^*(y) = \sup_{x \in X} (\langle y, x \rangle - f(x)) \quad \forall y \in X^*.$$

The double Fenchel transform  $f^{**}$  is the function on the space  $X$  defined by the expression

$$f^{**}(x) = \sup_{y \in X^*} (\langle y, x \rangle - f^*(y)) \quad \forall x \in X.$$

By Fenchel's theorem  $f^{**}(x) = \overline{\text{co}} f$  for arbitrary proper function  $f$ . This implies that for any proper function  $f$  its convex closure  $\overline{\text{co}} f$  coincides with the upper bound of the set of all affine continuous functions majorized by  $f$ .



**B. On coincidence of two definitions of the EoF.** Many questions related with the quantum entropy  $H(\rho)$  can be solved by using the following properties of this characteristic:

- nonnegativity;
- concavity;
- lower semicontinuity.

These properties follows from the substantially stronger property of the quantum entropy: the function  $H(\rho)$  is a pointwise limit of the increasing sequence of the nonnegative concave continuous and bounded <sup>8</sup> functions  $H(P_n \rho P_n) + \text{Tr } \rho P_n \log \text{Tr } \rho P_n$ , where  $P_n$  is arbitrary increasing sequence of finite rank projectors, converging to the unit operator in the strong operator topology [22].

The output entropy  $H_\Phi(\rho)$  of any quantum channel  $\Phi$  (in particular the partial trace channel) has all the above properties, which are used in proving many results of this paper. For example, the representation of the function  $H_\Phi(\rho)$  as a limit of increasing sequence of nonnegative concave continuous and bounded functions provides the proof of existence of  $\widehat{H}_\Phi$ -optimal measure supported by pure states (proposition 6).

The aim of this section is to show that the above properties of the output entropy  $H_\Phi(\rho)$  are not sufficient for proof of validity of equality (35) for all states.

Let  $\mathcal{E}$  be the class of all functions on  $\mathfrak{S}(\mathcal{H})$ , represented as a pointwise limit of increasing sequence of nonnegative concave continuous and bounded functions on  $\mathfrak{S}(\mathcal{H})$ .

This definition implies that all functions of the class  $\mathcal{E}$  are concave lower semicontinuous functions on  $\mathfrak{S}(\mathcal{H})$  with the range in  $[0, +\infty]$ .

PROPOSITION. *In  $\mathcal{E}$  there exists bounded function  $F$  such that*

$$\inf_{\sum_i \pi_i \rho_i = \rho_0} \sum_i \pi_i F(\rho_i) = 1 \quad \text{and} \quad \inf_{\pi \in \mathcal{P}(\rho_0)} \int_{\mathfrak{S}(\mathcal{H})} F(\rho) \pi(d\rho) = 0$$

for some state  $\rho_0$  in  $\mathfrak{S}(\mathcal{H})$ , where the infimum in the first expression is taken over all countable decompositions of the state  $\rho_0$ .

In the proof of this proposition the essential role is played by the following lemma, in which the indicator function of an arbitrary set of pure states is introduced.

LEMMA. *Let  $\mathcal{A}$  be an arbitrary set of pure states in  $\mathfrak{S}(\mathcal{H})$ . The function  $f_{\mathcal{A}}(\rho) = \inf_{\sigma \in \mathcal{A}} (1 - \text{Tr } \rho \sigma)$  is a continuous and concave function on  $\mathfrak{S}(\mathcal{H})$  such that  $0 \leq f(\rho) \leq 1$  and  $f^{-1}(0) = \overline{\mathcal{A}}$ .*

*Proof.* It is sufficient to prove convexity and continuity of the function  $1 - f(\rho) = g(\rho) = \sup_{\sigma \in \mathcal{A}} \text{Tr } \rho \sigma$ , since the other properties are easily verified. Convexity and lower semicontinuity of the function  $g(\rho)$  follows from its representation as an upper bound of the family  $\{\text{Tr } \rho \sigma\}_{\sigma \in \mathcal{A}}$  of continuous affine functions on  $\mathfrak{S}(\mathcal{H})$ .

Suppose the function  $g(\rho)$  is not upper semicontinuous. This means existence of such sequence of states  $\{\rho_n\}$ , converging to some state  $\rho_0$ , that

$$(36) \quad \lim_{n \rightarrow +\infty} g(\rho_n) > g(\rho_0).$$

Let  $\mathfrak{A} = \{|\varphi\rangle \in \mathcal{H} \mid |\varphi\rangle\langle\varphi| \in \mathcal{A}\}$  be the set of unit vectors in  $\mathcal{H}$  and  $\overline{\mathfrak{A}}$  be the closure

<sup>8</sup>Since the set  $\mathfrak{S}(\mathcal{H})$  is noncompact continuity does not imply boundedness.

of this set the weak topology in Hilbert space  $\mathcal{H}$ . By lemma 2 in [5, p. 284]<sup>9</sup> we have

$$(37) \quad g(\rho_0) = \sup_{\sigma \in \mathcal{A}} \text{Tr } \rho_0 \sigma = \sup_{\varphi \in \mathfrak{A}} \langle \varphi | \rho_0 | \varphi \rangle = \sup_{\varphi \in \overline{\mathfrak{A}}} \langle \varphi | \rho_0 | \varphi \rangle.$$

For any  $\varepsilon > 0$  and  $n$  there exists vector  $\varphi_n^\varepsilon$  in  $\mathfrak{A}$  such that  $\langle \varphi_n^\varepsilon | \rho_n | \varphi_n^\varepsilon \rangle > g(\rho_n) - \varepsilon$ . Since the unit ball in the space  $\mathcal{H}$  is compact in the weak topology, there exists subsequence  $\{\varphi_{n_k}^\varepsilon\}_k$  of the sequence  $\{\varphi_n^\varepsilon\}_n$ , weakly converging to some vector  $\varphi_0^\varepsilon \in \overline{\mathfrak{A}}$ . By lemma 2 in [5] mentioned above the sequence  $\{\langle \varphi_{n_k}^\varepsilon | \rho_0 | \varphi_{n_k}^\varepsilon \rangle\}_k$  converges to  $\langle \varphi_0^\varepsilon | \rho_0 | \varphi_0^\varepsilon \rangle$  as  $k \rightarrow +\infty$ . Thus by using the estimation  $|\langle \varphi_{n_k}^\varepsilon | \rho_{n_k} - \rho_0 | \varphi_{n_k}^\varepsilon \rangle| \leq \|\rho_{n_k} - \rho_0\|_1$  we obtain

$$\lim_{k \rightarrow +\infty} g(\rho_{n_k}) \leq \lim_{k \rightarrow +\infty} \langle \varphi_{n_k}^\varepsilon | \rho_{n_k} | \varphi_{n_k}^\varepsilon \rangle - \varepsilon = \langle \varphi_0^\varepsilon | \rho_0 | \varphi_0^\varepsilon \rangle - \varepsilon \leq g(\rho_0) - \varepsilon,$$

where the last inequality follows from (37). Since  $\varepsilon$  is arbitrary this inequality contradicts to (36). Lemma is proved.  $\blacksquare$

*Proof of Proposition.* Let  $\mathcal{A}_s$  be the set of all pure product states in the tensor product of two separable Hilbert spaces and  $\rho_0$  be the separable state, constructed in [3], such that any measure with the barycenter  $\rho_0$  has no atoms in  $\mathcal{A}_s$ . Let  $F$  be the characteristic function of the complement of the set  $\mathcal{A}_s$ . For each  $n = 1, 2, \dots$  the function  $F_n(\rho) = \sqrt[n]{f_{\mathcal{A}_s}(\rho)}$ , where  $f_{\mathcal{A}_s}(\rho)$  is the function from the above lemma, is continuous and concave. Hence, the pointwise limit  $F$  of the sequence  $\{F_n\}$  is a function of the class  $\mathcal{E}$ . By lemma 1 in [3] there exists (purely nonatomic) measure  $\pi_0$  supported by the set  $\mathcal{A}_s$  and having barycenter  $\rho_0$ . Thus  $\inf_{\pi \in \mathcal{P}_{\{\rho_0\}}} \int_{\mathfrak{S}(\mathcal{H})} F(\rho) \pi(d\rho) = 0$ . Since support of any atomic measure with the barycenter  $\rho_0$  does not intersect with  $\mathcal{A}_s$ , it easy to see that  $\inf_{\sum_i \pi_i \rho_i = \rho_0} \sum_i \pi_i F(\rho_i) = 1$ .

The author is grateful to A.S.Holevo for permanent help in preparing of this paper. The author is also grateful to M.Wolf for the useful remarks.

#### Список литературы

- [1] P.BILLINGSLEY, *Convergence of probability measures*, John Willey and Sons. Inc., New York-London-Sydney-Toronto.
- [2] O.BRATTELI, D.W.ROBINSON, *Operators algebras and quantum statistical mechanics*, Springer Verlag, New York-Heidelberg-Berlin, vol.I, 1979.
- [3] A.S.HOLEVO, M.E.SHIROKOV, R.F.WERNER, *On the notion of entanglement in Hilbert space*, Russian Math. Surveys, 2005, v. 60, No2.
- [4] A.D.JOFFE, W. M. TIKHOMIROV, *Theory of extremum problems*, AP, NY, 1979.
- [5] A.N.KOLMOGOROV, S.V.FOMIN, *Elements of function theory and functional analysis*, Moscow, Nauka, 1989 (In Russian).
- [6] E.S.POLOVINKIN, M.V. BALASHOV, *Elements of convex and strongly convex analysis*, M.: Physmatlit, 2004.
- [7] A.S.HOLEVO, *Introduction to quantum information theory*, MCCME, 2002 (in Russian).
- [8] A.S.HOLEVO, *Statistical structure of quantum theory*, Springer-Verlag, 2001.
- [9] A.S.HOLEVO, *Classical capacities of quantum channels with constrained inputs*, Theory Probab. Appl., 2003, V. 48, N. 2, p. 243–255.
- [10] A.S.HOLEVO, M.E.SHIROKOV, *Continuous ensembles and the  $\chi$ -capacity of infinite dimensional channels*, Theory Probab. Appl., 2005, V. 50, N. 1, p. 86–98.
- [11] M.E.SHIROKOV, *Entropic characteristics of subsets of states*, I, Izvestiya RAN: Ser. Mat., 2006, V. 70, No6, p. 1265–1292.

<sup>9</sup>This lemma asserts that for arbitrary sequence of vectors  $\{x_n\}$  in Hilbert space, weakly converging to some vector  $x_0$ , and for arbitrary compact operator  $A$  there exists  $\lim_{n \rightarrow +\infty} \langle x_n | A | x_n \rangle = \langle x_0 | A | x_0 \rangle$ .

- [12] E. ALFSEN, *Compact Convex Sets and Boundary Integrals*, Springer-Verlag, New York, Heidelberg, 1971, 210 p.
- [13] K. M. R. AUDENAERT AND S. L. BRAUNSTEIN, *On strong subadditivity of the entanglement of formation*, *Comm. Math. Phys.*, 2004, v. 246, No3, pp. 443–452.
- [14] R. D. BOURGIN, *Partial orderings for integral representations on convex sets with the Radon–Nikodým property*, *Pacific J. Math.*, 1979, v. 81, No1, pp. 29–44.
- [15] M. J. DONALD, *Further results on the relative entropy*, *Math. Proc. Cambridge Philos. Soc.*, 1987, v. 101, No2, pp. 363–373.
- [16] G. F. DELL’ANTONIO, *On the limits of sequences of normal states*, *Comm. Pure Appl. Math.*, 1967, v. 20, pp. 413–429.
- [17] C. H. BENNETT, D. P. DIVINCENZO, J. A. SMOLIN, AND W. K. WOOTTERS, *Mixed-state entanglement and quantum error correction*, *Phys. Rev. A* (3), 1996, v. 54, No5, pp. 3824–3851.
- [18] E. B. DAVIES, *Information and quantum measurement*, *IEEE Trans. Inform. Theory*, 1978, v. 24, No5, pp. 596–599.
- [19] J. EISERT, C. SIMON, AND M. B. PLENIO, *On the quantification of entanglement in infinite-dimensional quantum systems*, *J. Phys. A*, 2002, v. 35, No17, pp. 3911–3923.
- [20] P. HARREMOËS, *Information topologies with applications*, *Entropy, Search, Complexity*, Ed. by I. Császár, G. O. H. Katona, and G. Tardos. Springer-Verlag, Berlin, 2007, pp. 113–150. (Bolyai Soc. Math. Stud., v. 16.)
- [21] A. S. HOLEVO AND M. E. SHIROKOV, *On Shor’s channel extension and constrained channels*, *Comm. Math. Phys.*, 2004, v. 249, No2, pp. 417–430.
- [22] G. LINDBLAD, *Expectation and entropy inequalities for finite quantum systems*, *Comm. Math. Phys.*, 1974, v. 39, No2, pp. 111–119.
- [23] G. LINDBLAD, *Completely positive maps and entropy inequalities*, *Comm. Math. Phys.*, 1975, v. 40, No2, pp. 147–151.
- [24] A. W. MAJEWSKI, *On entanglement of formation*, *J. Phys. A*, 2002, v. 35, No1, pp. 123–134.
- [25] M. A. NIELSEN AND I. L. CHUANG, *Quantum Computation and Quantum Information*, Cambridge Univ. Press, Cambridge, 2000, 676 p.
- [26] M. A. NIELSEN, *Continuity bounds for entanglement*, <http://arxiv.org: quant-ph/9908086>, 1999.
- [27] M. OHYA AND D. PETZ, *Quantum Entropy and Its Use*, Springer-Verlag, Berlin, 1993, 335 p.
- [28] K. R. PARTHASARATHY, *Probability Measures on Metric Spaces*, Academic Press, New York, London, 1967, 276 p.
- [29] B. SCHUMACHER AND M. D. WESTMORELAND, *Optimal signal ensembles*, <http://arxiv.org: quant-ph/9912122>, 1999.
- [30] M. E. SHIROKOV, *On the additivity conjecture for channels with arbitrary constraints*, <http://arxiv.org: quant-ph/0308168>, 2003.
- [31] M. E. SHIROKOV, *The Holevo capacity of infinite dimensional channels and the additivity problem*, *Comm. Math. Phys.*, 2006, v. 262, No1, pp. 137–159.
- [32] M. E. SHIROKOV, *On channels with finite Holevo capacity*, <http://arxiv.org: quant-ph/0602073>, 2006.
- [33] B. SIMON, *Convergence theorem for entropy*. Appendix in LIEB E. H. AND RUSKAI M. B. *Proof of the strong subadditivity of quantum-mechanical entropy*, *J. Math. Phys.*, 1973, v. 14, pp. 1938–1941.
- [34] A. WEHRL, *General properties of entropy*, *Rev. Modern Phys.*, 1978, v. 50, No2, pp. 221–260.
- [35] A. UHLMANN, *Entropy and optimal decompositions of states relative to a maximal commutative subalgebra*, <http://arxiv.org: quant-ph/9704017>, 1997.