

# Continuity condition for concave functions on convex $\mu$ -compact sets and its applications in quantum physics

M.E. Shirokov\*

Steklov Mathematical Institute, Moscow, Russia

## Abstract

A method of proving local continuity of concave functions on convex set possessing the  $\mu$ -compactness property is presented. This method is based on a special approximation of these functions.

The class of  $\mu$ -compact sets can be considered as a natural extension of the class of compact metrizable subsets of locally convex spaces, to which particular results well known for compact sets can be generalized.

Applications of the obtained continuity conditions to analysis of different entropic characteristics of quantum systems and channels are considered.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Basic notations</b>	<b>2</b>
<b>3</b>	<b>Some implications of <math>\mu</math>-compactness and stability of a convex set</b>	<b>5</b>
<b>4</b>	<b>Special approximation of concave functions</b>	<b>9</b>
<b>5</b>	<b>Continuity conditions</b>	<b>10</b>
5.1	The case $\mathcal{A} = \sigma\text{-co}(\text{cl}(\text{extr}\mathcal{A}))$ . . . . .	10
5.2	Possible generalizations . . . . .	13
<b>6</b>	<b>Applications in quantum physics</b>	<b>14</b>
6.1	The von Neumann entropy . . . . .	16
6.2	The quantum mutual information . . . . .	17
6.3	The output von Neumann entropy and the $\chi$ -function of a quantum channel	19
<b>7</b>	<b>Appendix</b>	<b>22</b>

---

\*e-mail:msh@mi.ras.ru

# 1 Introduction

A problem of proving local continuity of a given concave (convex) function defined on a given convex set arises naturally in different fields of mathematics. For example, in mathematical physics this problem appears in analysis of entropy-type functions on a set of states of some physical system. In some cases this problem can be solved by using general results of convex analysis, but sometimes it is difficult to apply them to a given function defined on a convex set not satisfying particular requirements (compactness, existence of inner points, etc.)

In this paper we consider local continuity conditions for concave functions on  $\mu$ -compact convex sets based on a special approximation of these functions. The class of  $\mu$ -compact sets (see Definition 1 in Section 2) can be considered as a natural extension of the class of compact metrizable subsets of locally convex spaces, to which particular results well known for compact sets can be generalized [24]. This class contains all compact sets as well as many noncompact sets widely used in applications. The simplest examples of noncompact  $\mu$ -compact convex sets are the positive part of the unit ball of the Banach space  $\ell_1$  and its closed subset consisting of all countable probability distributions. Other examples and simple criteria of the  $\mu$ -compactness property can be found in [24].

For applications in quantum physics it is essential that the convex set of positive operators in a separable Hilbert space with unit trace, generally called quantum states, is  $\mu$ -compact. In fact, it is necessary to explore continuity properties of several entropic characteristics of quantum states, in particular, of the von Neumann entropy, that provides a basic impetus to find universal method of proving local continuity of these characteristics. In [27] this method was developed by using some special properties of the set of quantum states. In this paper we show that it can be generalized to the class of  $\mu$ -compact convex sets by using slightly different argumentation.

The paper is organized as follows. In Section 2 notations and basic results used in the subsequent sections are presented. In Section 3 we consider several properties of a convex set following from its  $\mu$ -compactness and stability (see Definition 2 in Section 2). Section 4 is devoted to a special approximation technic for concave functions. In Section 5 the continuity conditions based on this technic are presented. Applications to quantum physics extending the results of [27] are considered in Section 6.

## 2 Basic notations

In what follows  $\mathcal{A}$  is a bounded convex complete separable metrizable subset of some locally convex space.<sup>1</sup> The set of extreme points of the set  $\mathcal{A}$  will be denoted  $\text{extr}(\mathcal{A})$ .

Let  $\text{cl}(\mathcal{B})$ ,  $\text{co}(\mathcal{B})$ ,  $\sigma\text{-co}(\mathcal{B})$  and  $\overline{\text{co}}(\mathcal{B})$  be respectively the closure, the convex hull, the  $\sigma$ -convex hull<sup>2</sup> and the convex closure of a subset  $\mathcal{B} \subseteq \mathcal{A}$  [17, 25].

---

<sup>1</sup>This means that the topology on the set  $\mathcal{A}$  is defined by a countable subset of the family of seminorms, generating the topology of the entire locally convex space, and this set is separable and complete in the metric generated by this subset of seminorms.

<sup>2</sup> $\sigma\text{-co}(\mathcal{B})$  is the set of all countable convex combinations of points in  $\mathcal{B}$ .

For an arbitrary closed subset  $\mathcal{B} \subseteq \mathcal{A}$  denote by  $C(\mathcal{B})$  the set of all continuous bounded functions on  $\mathcal{B}$ , denote by  $M(\mathcal{B})$  and  $M^a(\mathcal{B})$  respectively the set of all Borel probability measures on  $\mathcal{B}$  and its subset consisting of atomic measures. We always assume that the set  $M(\mathcal{B})$  and arbitrary its subsets are endowed with the weak convergence topology [7, 23].

With an arbitrary measure  $\mu \in M(\mathcal{B})$  we associate its barycenter (average)  $\mathbf{b}(\mu) \in \overline{\text{co}}(\mathcal{B})$ , which is defined by the Pettis integral (see [3, 29])

$$\mathbf{b}(\mu) = \int_{\mathcal{B}} x \mu(dx). \quad (1)$$

If  $\mu$  is a measure in  $M^a(\mathcal{B})$  "consisting" of atoms  $\{x_i\}$  with the corresponding weights  $\{\pi_i\}$  then  $\mathbf{b}(\mu) = \sum_i \pi_i x_i$ . The above measure will be denoted  $\sum_i \pi_i \delta(x_i)$  or, briefly,  $\{\pi_i, x_i\}$ .

For a given Borel function  $f$  on a closed subset  $\mathcal{B} \subseteq \mathcal{A}$  consider the functional

$$M(\mathcal{B}) \ni \mu \mapsto \mathbf{f}(\mu) = \int_{\mathcal{B}} f(x) \mu(dx). \quad (2)$$

It is easy to show that this functional is lower semicontinuous (correspondingly, upper semicontinuous) provided the function  $f$  is lower semicontinuous and lower bounded (correspondingly, upper semicontinuous and upper bounded) on the set  $\mathcal{B}$  [7].

For arbitrary  $x \in \overline{\text{co}}(\mathcal{B})$  let  $M_x(\mathcal{B})$  and  $M_x^a(\mathcal{B})$  be respectively convex subsets of  $M(\mathcal{B})$  and of  $M^a(\mathcal{B})$  consisting of such measures  $\mu$  that  $\mathbf{b}(\mu) = x$ .

The barycenter map

$$M(\mathcal{A}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A} \quad (3)$$

is continuous (this can be shown easily by applying Prokhorov's theorem [23, Ch.II, Th.6.7]). Hence the image of any compact subset of  $M(\mathcal{A})$  under this map is a compact subset of  $\mathcal{A}$ . The  $\mu$ -compact sets are defined by the converse requirement [24].

**Definition 1.** A set  $\mathcal{A}$  is called  $\mu$ -compact if the preimage of any compact subset of  $\mathcal{A}$  under barycenter map (3) is a compact subset of  $M(\mathcal{A})$ .

Any compact set is  $\mu$ -compact, since compactness of  $\mathcal{A}$  implies compactness of  $M(\mathcal{A})$  [23]. Properties of  $\mu$ -compact sets are studied in detail in [24], where  $\mu$ -compactness of several noncompact sets widely used in applications has been proved (for example, of the set of all Borel probability measures on an arbitrary complete separable metric space endowed with the weak convergence topology and of the set of quantum states – density operators in a separable Hilbert space).

The  $\mu$ -compactness property of a convex set is not purely topological but reflects a special relation between the topology and the convex structure of this set.

Another relation between the topology and the convex structure of a convex set is expressed by the notion of (convex) stability [22].

**Definition 2.** A set  $\mathcal{A}$  is called *stable* if the map  $\mathcal{A} \times \mathcal{A} \ni (x, y) \mapsto \frac{x + y}{2} \in \mathcal{A}$  is open.

The notion of stability of a convex subset of a linear topological space appeared at the end of 1970's as a result of study of convex compact sets, which leded in particular to proving equivalence of the following properties a convex compact set  $\mathcal{A}$ :

- (i) *the set  $\mathcal{A}$  is stable;*
- (ii) *the map  $M(\mathcal{A}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$  is open;*
- (iii) *the map  $M(\overline{\text{extr } \mathcal{A}}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$  is open;*
- (iv) *the convex hull<sup>3</sup> of an arbitrary continuous function on  $\mathcal{A}$  is continuous;*
- (v) *the convex hull of an arbitrary concave continuous function on  $\mathcal{A}$  is continuous.*

Essential parts of the above assertion was obtained by Vesterstrom [30], its complete version was proved by O'Brien [9]. This assertion (called the Vesterstrom-O'Brien theorem in what follows) does not hold for noncompact convex sets in general, but it can be extended to convex  $\mu$ -compact sets [24, Theorem 1].

In  $\mathbb{R}^2$  stability holds for an arbitrary convex compact set, in  $\mathbb{R}^3$  it is equivalent to closedness of the set of extreme points of a convex compact set while in  $\mathbb{R}^n, n > 3$ , it is stronger than the last property [9]. A full characterization of the stability property in finite dimensions is obtained in [22]. In infinite dimensions stability is proved for the unit ball in some Banach spaces and for the positive part of the unit ball in Banach lattices in which the unit ball is stable [10].

The simplest example of a noncompact  $\mu$ -compact convex stable set is the set  $\mathfrak{P}_{+\infty}$  of all probability distributions with countable number of outcomes (considered as a subset of the Banach space  $\ell_1$ ). This is a partial case of the more general example – the convex set of all Borel probability measures on any complete separable metric space endowed with the weak convergence topology. The  $\mu$ -compactness and stability of this set are established respectively in [24, Corollary 4] and in [12, Theorem 2.4].

We will use the following two strengthened versions of the notion of a concave function.

A semibounded (upper or lower bounded) function  $f$  on a convex set  $\mathcal{A}$  is called  $\sigma$ -concave if the discrete Jensen's inequality

$$f(\mathbf{b}(\{\pi_i, x_i\})) \geq \sum_i \pi_i f(x_i)$$

holds for an arbitrary measure  $\{\pi_i, x_i\}$  in  $M^a(\mathcal{A})$ .

A semibounded universally measurable<sup>4</sup> function  $f$  on a convex set  $\mathcal{A}$  is called  $\mu$ -concave if the integral Jensen's inequality

$$f(\mathbf{b}(\mu)) \geq \int_{\mathcal{A}} f(x)\mu(dx)$$

---

<sup>3</sup>The convex hull of a function is the maximal convex function majorized by this function [17].

<sup>4</sup>This means that the function  $f$  is measurable with respect to any measure in  $M(\mathcal{A})$ .

holds for an arbitrary measure  $\mu$  in  $M(\mathcal{A})$ .

$\sigma$ -convexity and  $\mu$ -convexity of a function  $f$  are naturally defined via the above notions applied to the function  $-f$ .

Examples of semibounded functions, which are convex but not  $\sigma$ -convex or  $\sigma$ -convex but not  $\mu$ -convex, are considered in [28, Section 3].

The following lemma contains sufficient conditions for  $\sigma$ -concavity and  $\mu$ -concavity of a concave function, which can be proved easily (see the Appendix in [28]).

**Lemma 1.** *Let  $f$  be a concave function on a convex set  $\mathcal{A}$ .*

- A) *If  $f$  is lower bounded then  $f$  is  $\sigma$ -concave.*
- B) *If  $f$  is either lower semicontinuous and lower bounded or upper semicontinuous then  $f$  is  $\mu$ -concave.*

**Remark 1.** Assuming that the metric  $\mathbf{d}(\cdot, \cdot)$  on the set  $\mathcal{A}$  is defined as follows

$$\mathbf{d}(x, y) = \sum_{k=1}^{+\infty} 2^{-k} \frac{\|x - y\|_k}{1 + \|x - y\|_k}, \quad x, y \in \mathcal{A},$$

where  $\{\|\cdot\|_k\}_{k=1}^{+\infty}$  is the countable family of seminorms generating the topology on this set, it is easy to obtain the following estimation

$$\mathbf{d}(\alpha x + (1 - \alpha)y, \alpha'x' + (1 - \alpha')y') \leq 2\delta + C_{x,y}(\varepsilon)$$

valid for any  $x, y, x', y'$  in  $\mathcal{A}$  and any  $\alpha, \alpha'$  in  $[0, 1]$  such that  $\mathbf{d}(x, x') < \delta$ ,  $\mathbf{d}(y, y') < \delta$  and  $|\alpha - \alpha'| < \varepsilon$ , where  $C_{x,y}(\varepsilon) = \sum_{k=1}^{+\infty} 2^{-k} \frac{\varepsilon \|x - y\|_k}{1 + \varepsilon \|x - y\|_k}$  is a function such that  $\lim_{\varepsilon \rightarrow +0} C_{x,y}(\varepsilon) = 0$ .

**Note:** In what follows continuity of a function  $f$  on a subset  $\mathcal{B} \subset \mathcal{A}$  means continuity of the restriction  $f|_{\mathcal{B}}$  of the function  $f$  to the subset  $\mathcal{B}$ , which implies finiteness of this restriction (in contrast to lower or upper semicontinuity).

### 3 Some implications of $\mu$ -compactness and stability of a convex set

In this section we consider auxiliary results used in the main part of the paper.

We begin with several simple lemmas.

**Lemma 2.** *Let  $\mathcal{B}$  be a closed subset of a convex  $\mu$ -compact set  $\mathcal{A}$ . Then for arbitrary  $x_0$  in  $\overline{\text{co}}(\mathcal{B})$  there exists a measure  $\mu_0$  in  $M(\mathcal{B})$  such that  $x_0 = \mathbf{b}(\mu_0)$ .*

**Proof.** Let  $x_0 \in \overline{\text{co}}(\mathcal{B})$  and  $\{x_n\} \subset \text{co}(\mathcal{B})$  be a sequence converging to  $x_0$ . For each  $n \in \mathbb{N}$  there exists a measure  $\mu_n \in M(\mathcal{B})$  with finite support such that  $x_n = \mathbf{b}(\mu_n)$ . By  $\mu$ -compactness of the set  $\mathcal{A}$  the sequence  $\{\mu_n\}$  has a partial limit  $\mu_0 \in M(\mathcal{B})$ . Continuity of the map  $\mu \mapsto \mathbf{b}(\mu)$  implies  $\mathbf{b}(\mu_0) = x_0$ .  $\square$

**Lemma 3.** *Let  $\mathcal{A}$  be a convex  $\mu$ -compact set such that the set  $\text{extr}\mathcal{A}$  is closed and  $\mathcal{A} = \sigma\text{-co}(\text{extr}\mathcal{A})$ . Then an arbitrary measure  $\mu_0$  in  $M(\text{extr}\mathcal{A})$  can be approximated by a sequence  $\{\mu_n\}$  of measures in  $M^a(\text{extr}\mathcal{A})$  such that  $\mathbf{b}(\mu_n) = \mathbf{b}(\mu_0)$  for all  $n$ .*

**Proof.** Consider the Choquet ordering on the set  $M(\mathcal{A})$ . We say that  $\mu \succ \nu$  if and only if

$$\int_{\mathcal{A}} f(x)\mu(dx) \geq \int_{\mathcal{A}} f(x)\nu(dx)$$

for any convex continuous bounded function  $f$  on the set  $\mathcal{A}$  [11].

For a given measure  $\mu_0$  in  $M(\text{extr}\mathcal{A})$  it is easy to construct a sequence  $\{\mu_n\}$  of measures in  $M(\mathcal{A})$  with finite support converging to the measure  $\mu_0$  such that  $\mathbf{b}(\mu_n) = \mathbf{b}(\mu_0)$  for all  $n$ . Decomposing each atom of the measure  $\mu_n$  into convex combination of extreme points we obtain the measure  $\hat{\mu}_n$  in  $M^a(\text{extr}\mathcal{A})$  with the same barycenter. It is easy to see that  $\hat{\mu}_n \succ \mu_n$ . By  $\mu$ -compactness of the set  $\mathcal{A}$  the sequence  $\{\hat{\mu}_n\}_{n>0}$  is relatively compact. This implies existence of subsequence  $\{\hat{\mu}_{n_k}\}$  converging to a measure  $\hat{\mu}_0$  in  $M(\text{extr}\mathcal{A})$ . Since  $\hat{\mu}_{n_k} \succ \mu_{n_k}$  for all  $k$ , the definition of the weak convergence implies  $\hat{\mu}_0 \succ \mu_0$  and hence  $\hat{\mu}_0 = \mu_0$  by maximality of the measure  $\mu_0$  with respect to the Choquet ordering (which follows from coincidence of this ordering with the dilation ordering [11]).  $\square$

**Lemma 4.** *Let  $\mathcal{A}$  be a convex  $\mu$ -compact set and  $\{\{\pi_i^n, x_i^n\}_{i=1}^m\}_n$  be a sequence of measures in  $M^a(\mathcal{A})$  having  $m < +\infty$  atoms such that the sequence  $\{\sum_{i=1}^m \pi_i^n x_i^n\}_n$  of their barycenters converges to a point  $x_0 \in \mathcal{A}$ . There exists a subsequence  $\{\{\pi_i^{n_k}, x_i^{n_k}\}_{i=1}^m\}_k$  converging to a particular measure<sup>5</sup>  $\{\pi_i^0, x_i^0\}_{i=1}^m$  with the barycenter  $x_0$  in the following sense*

$$\lim_{k \rightarrow +\infty} \pi_i^{n_k} = \pi_i^0 \quad \text{and} \quad \pi_i^0 > 0 \Rightarrow \lim_{k \rightarrow +\infty} x_i^{n_k} = x_i^0, \quad i = \overline{1, m}.$$

**Proof.** It is sufficient to note that  $\mu$ -compactness of the set  $\mathcal{A}$  implies relative compactness of the sequence  $\{\{\pi_i^n, x_i^n\}_{i=1}^m\}_n$  and that the set of measures having  $m$  atoms is a closed subset of  $M(\mathcal{A})$ .  $\square$

Let  $\mathfrak{P}_n$  be the set of all probability distributions with  $n \leq +\infty$  outcomes.

**Lemma 5.** *Let  $\mathcal{A}_1$  be a closed subset of a convex  $\mu$ -compact set  $\mathcal{A}$ .*

A) *The set*

$$\mathcal{A}_k = \left\{ \sum_{i=1}^k \pi_i x_i \mid \{\pi_i\} \in \mathfrak{P}_k, \{x_i\} \subset \mathcal{A}_1 \right\} \quad (4)$$

*is closed for each  $k \in \mathbb{N}$ .*

B) *Let  $f$  be a concave nonnegative function on the set  $\mathcal{A}$ , which takes a finite value at least at one point in  $\mathcal{A}_1$ . If this function is upper continuous on the set  $\mathcal{A}_k$  defined by (4) for each  $k$  then it is bounded on the set  $\mathcal{A}_k$  for each  $k$ .*

---

<sup>5</sup>We do not assert that  $x_i^0 \neq x_j^0$  for all  $i \neq j$ .

**Proof.** A) This assertion directly follows from Lemma 4.

B) Suppose there exists a sequence  $\{x_n\} \subset \mathcal{A}_k$  such that  $\lim_{n \rightarrow +\infty} f(x_n) = +\infty$ . Let  $y_0$  be a point in  $\mathcal{A}_1$  with finite  $f(y_0)$ . Consider the sequence  $\{\lambda_n x_n + (1 - \lambda_n)y_0\} \subset \mathcal{A}_{k+1}$ , where  $\lambda_n = 1/f(x_n)$ . This sequence converges to the point  $y_0$  (since the set  $\mathcal{A}$  is bounded), but concavity of the function  $f$  implies

$$\liminf_{n \rightarrow +\infty} f(\lambda_n x_n + (1 - \lambda_n)y_0) \geq \liminf_{n \rightarrow +\infty} (\lambda_n f(x_n) + (1 - \lambda_n)f(y_0)) = 1 + f(y_0),$$

contradicting upper semicontinuity of the function  $f$  on the set  $\mathcal{A}_{k+1}$ .  $\square$

An essential property of  $\mu$ -compact sets is presented in the following proposition.

**Proposition 1.** *Let  $\mathcal{A}$  be a convex  $\mu$ -compact set and let  $f$  be an upper semicontinuous upper bounded function on a closed subset  $\mathcal{B} \subset \mathcal{A}$ . Then the function*

$$\hat{f}_{\mathcal{B}}^{\mu}(x) = \sup_{\mu \in M_x(\mathcal{B})} \int_{\mathcal{B}} f(y) \mu(dy) \quad (5)$$

is upper semicontinuous and  $\mu$ -concave on the set  $\overline{\text{co}}(\mathcal{B})$ . For arbitrary  $x \in \overline{\text{co}}(\mathcal{B})$  the supremum in the definition of the value  $\hat{f}_{\mathcal{B}}^{\mu}(x)$  is achieved at a particular measure in  $M_x(\mathcal{B})$ .

This property provides generalization of several results well known for compact convex sets to  $\mu$ -compact convex sets [24, Proposition 6, Corollary 2]. It becomes no valid after slight relaxing of the  $\mu$ -compactness assumption to pointwise  $\mu$ -compactness [24, Proposition 7]. The proof of Proposition 1 is placed in the Appendix.

Another important technical tool is presented in the following proposition.

**Proposition 2.** *Let  $\mathcal{A}$  be a convex  $\mu$ -compact<sup>6</sup> set and let  $f$  be a lower semicontinuous lower bounded function on a closed subset  $\mathcal{B} \subseteq \mathcal{A}$ .*

A) *If the map  $M(\mathcal{B}) \ni \mu \mapsto \mathbf{b}(\mu) \in \overline{\text{co}}(\mathcal{B})$  is open then the function  $\hat{f}_{\mathcal{B}}^{\mu}$  defined by (5) is lower semicontinuous and  $\mu$ -concave on the set  $\overline{\text{co}}(\mathcal{B})$ .*

B) *If the map  $M^{\alpha}(\mathcal{B}) \ni \mu \mapsto \mathbf{b}(\mu) \in \sigma\text{-co}(\mathcal{B})$  is open then the  $\sigma$ -concave function*

$$\hat{f}_{\mathcal{B}}^{\sigma}(x) = \sup_{\mu \in M_x^{\alpha}(\mathcal{B})} \int_{\mathcal{B}} f(y) \mu(dy) = \sup_{\{\pi_i, x_i\} \in M_x^{\alpha}(\mathcal{B})} \sum_i \pi_i f(x_i)$$

is lower semicontinuous on the set  $\sigma\text{-co}(\mathcal{B})$ . If, in addition,  $\sigma\text{-co}(\mathcal{B}) = \overline{\text{co}}(\mathcal{B})$  then the function  $\hat{f}_{\mathcal{B}}^{\sigma}$  coincides with the function  $\hat{f}_{\mathcal{B}}^{\mu}$  defined by (5).

The proof of Proposition 2 is placed in the Appendix.

**Remark 2.** If  $f$  is bounded and upper semicontinuous function on a closed subset  $\mathcal{B}$  of a convex  $\mu$ -compact set  $\mathcal{A}$  such that  $\sigma\text{-co}(\mathcal{B}) = \overline{\text{co}}(\mathcal{B})$  then the above defined functions  $\hat{f}_{\mathcal{B}}^{\sigma}$

---

<sup>6</sup>The  $\mu$ -compactness assumption is used only to guarantee  $\mathbf{b}(M(\mathcal{B})) = \overline{\text{co}}(\mathcal{B})$  by means of Lemma 2.

and  $\hat{f}_{\mathcal{B}}^{\mu}$  do not coincide in general (see the example in [28, Remark 9]). Thus the assertion of Propositions 1 does not hold for the function  $\hat{f}_{\mathcal{B}}^{\sigma}$  (since  $\mu$ -concavity of  $\hat{f}_{\mathcal{B}}^{\sigma}$  implies  $\hat{f}_{\mathcal{B}}^{\sigma} = \hat{f}_{\mathcal{B}}^{\mu}$ ).

Propositions 1 and 2 have the obvious corollary.

**Corollary 1.** *Let  $\mathcal{B}$  be a closed subset of a convex  $\mu$ -compact set  $\mathcal{A}$ .*

- A) *If  $\mathcal{A} = \overline{\text{co}}(\mathcal{B})$  and the map  $M(\mathcal{B}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$  is open then  $\hat{f}_{\mathcal{B}}^{\mu} \in C(\mathcal{A})$  for any  $f \in C(\mathcal{B})$ .*
- B) *If  $\mathcal{A} = \sigma\text{-co}(\mathcal{B})$  and the map  $M^a(\mathcal{B}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$  is open then  $\hat{f}_{\mathcal{B}}^{\sigma} = \hat{f}_{\mathcal{B}}^{\mu} \in C(\mathcal{A})$  for any  $f \in C(\mathcal{B})$ .*

If  $\mathcal{A}$  is a stable convex  $\mu$ -compact set then the set  $\text{extr}\mathcal{A}$  is closed and the generalized Vesterstrom-O'Brien theorem ([24, Theorem 1]) implies openness of the surjective map  $M(\text{extr}\mathcal{A}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$ , hence Corollary 1A shows that an arbitrary function  $f$  in  $C(\text{extr}\mathcal{A})$  has continuous bounded concave extension  $\hat{f}_{\text{extr}\mathcal{A}}^{\mu}$  to the set  $\mathcal{A}$ . This property does not hold in general for stable convex sets, which are not  $\mu$ -compact (see Example 1 in [24]).

Corollary 1B plays an essential role in this paper due to the following observation.

**Proposition 3.** *Let  $\mathcal{A}_1$  be a closed subset of a stable convex  $\mu$ -compact set  $\mathcal{A}$  such that  $\mathcal{A} = \sigma\text{-co}(\mathcal{A}_1)$  and  $\{\mathcal{A}_k\}$  be the family of subsets defined by (4). If the map*

$$M^a(\mathcal{A}_k) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A} \tag{6}$$

*is open for  $k = 1$  then this map is open for all  $k \in \mathbb{N}$ .*

The proof of Proposition 3 is placed in the Appendix.

By the generalized Vesterstrom-O'Brien theorem stability of a convex  $\mu$ -compact set  $\mathcal{A}$  is equivalent to openness of the map  $M(\text{extr}\mathcal{A}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$ . By Lemma 3 the last property implies openness of the map  $M^a(\text{extr}\mathcal{A}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$ . Hence we obtain from Proposition 3 the following assertion.

**Corollary 2.** *Let  $\mathcal{A}$  be a stable convex  $\mu$ -compact set such that  $\mathcal{A} = \sigma\text{-co}(\text{extr}\mathcal{A})$  and  $\{\mathcal{A}_k\}$  be the family of subsets defined by (4) with  $\mathcal{A}_1 = \text{extr}\mathcal{A}$ . Then map (6) is open for all  $k \in \mathbb{N}$ .*

**Remark 3.** If  $\mathcal{A}$  is the stable convex  $\mu$ -compact set of quantum states (see Section 6) and  $\mathcal{A}_1 = \text{extr}\mathcal{A}$  then openness of the maps (6) and

$$M(\mathcal{A}_k) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A} \tag{7}$$

are proved in [27] by using special structure of this set and called strong stability property. By Corollary 3 to prove that this strong stability property follows from stability it suffices to show that openness of map (7) follows from openness of map (6). In [27] this is made by proving density of the set  $M_x^a(\mathcal{A}_k)$  in  $M_x(\mathcal{A}_k)$  for all  $x \in \mathcal{A}$ .

**Question 1.** Let  $\mathcal{A}$  be a convex  $\mu$ -compact set and  $\{\mathcal{A}_k\}$  be the family of subsets defined by (4) with  $\mathcal{A}_1 = \text{extr}\mathcal{A}$ . Does stability of the set  $\mathcal{A}$  imply openness of map (7) for all  $k$ ?

A positive answer on this question can be used to strengthen Theorem 2 in Section 5 (see Remark 7 after this theorem).



## 4 Special approximation of concave functions

Throughout this section we will assume that  $f$  is a concave nonnegative function on a convex  $\mu$ -compact set  $\mathcal{A}$  having universally measurable restrictions to subsets of the family  $\{\mathcal{A}_k\}$  defined by (4) with  $\mathcal{A}_1 = \text{cl}(\text{extr}\mathcal{A})$ . By Lemma 5 this family consists of closed subsets. Possible generalizations are mentioned in Remark 4 at the end of this section.

Since for arbitrary  $x$  in  $\mathcal{A}$  the set  $M_x(\mathcal{A}_1)$  is not empty by Proposition 5 in [24], for given natural  $k$  we can consider the concave nonnegative function

$$\mathcal{A} \ni x \mapsto \hat{f}_k^\mu(x) = \sup_{\mu \in M_x(\mathcal{A}_k)} \int_{\mathcal{A}_k} f(y) \mu(dy). \quad (8)$$

If the function  $f$  is  $\mu$ -concave on the set  $\mathcal{A}$  then

$$\hat{f}_k^\mu \leq f \quad \text{and} \quad \hat{f}_k^\mu|_{\mathcal{A}_k} = f|_{\mathcal{A}_k}, \quad (9)$$

hence the function  $\hat{f}_k^\mu$  can be considered as a concave extension of the function  $f|_{\mathcal{A}_k}$  to the set  $\mathcal{A}$ . If the function  $f$  has upper semicontinuous bounded restriction to the set  $\mathcal{A}_k$  then the bounded function  $\hat{f}_k^\mu$  is upper semicontinuous and  $\mu$ -concave on the set  $\mathcal{A}$  by Proposition 1. Hence in this case the function  $\hat{f}_k^\mu$  is the minimal  $\mu$ -concave extension of the function  $f|_{\mathcal{A}_k}$  to the set  $\mathcal{A}$ .

The sequence  $\{\hat{f}_k^\mu\}$  is nondecreasing and its pointwise limit  $\hat{f}_*^\mu \doteq \sup_k \hat{f}_k^\mu$  is a concave function on  $\mathcal{A}$ . If the function  $f$  is  $\mu$ -concave then (9) implies

$$\hat{f}_*^\mu \leq f \quad \text{and} \quad \hat{f}_*^\mu|_{\mathcal{A}_*} = f|_{\mathcal{A}_*}, \quad \text{where} \quad \mathcal{A}_* = \bigcup_{k=1}^{+\infty} \mathcal{A}_k. \quad (10)$$

**Question 2.** Under what conditions do the functions  $\hat{f}_*^\mu$  and  $f$  coincide?

A partial answer on this question can be obtained in the case  $\mathcal{A} = \sigma\text{-co}(\mathcal{A}_1)$ .

In this case for given natural  $k$  one can consider the  $\sigma$ -concave nonnegative function

$$\mathcal{A} \ni x \mapsto \hat{f}_k^\sigma(x) = \sup_{\{\pi_i, x_i\} \in M_x^\sigma(\mathcal{A}_k)} \sum_i \pi_i f(x_i). \quad (11)$$

By the construction  $\hat{f}_k^\sigma \leq \hat{f}_k^\mu$ . Since the function  $f$  is  $\sigma$ -concave on the set  $\mathcal{A}$  by Lemma 1, we have

$$\hat{f}_k^\sigma \leq f \quad \text{and} \quad \hat{f}_k^\sigma|_{\mathcal{A}_k} = f|_{\mathcal{A}_k}. \quad (12)$$

Hence the function  $\hat{f}_k^\sigma$  is the minimal  $\sigma$ -concave extension of the function  $f|_{\mathcal{A}_k}$  to the set  $\mathcal{A}$ .

The sequence  $\{\hat{f}_k^\sigma\}$  is nondecreasing and its pointwise limit  $\hat{f}_*^\sigma \doteq \sup_k \hat{f}_k^\sigma$  is a concave function on  $\mathcal{A}$  such that  $\hat{f}_*^\sigma \leq \hat{f}_*^\mu$ . It follows from (12) that relations (10) hold with  $\hat{f}_*^\sigma$  instead of  $\hat{f}_*^\mu$ .

**Proposition 4.** *If  $\mathcal{A} = \sigma\text{-co}(\mathcal{A}_1)$  and the function  $f$  is lower semicontinuous then*

$$\hat{f}_*^\mu = \hat{f}_*^\sigma = f.$$

**Proof.** By Lemma 1 the function  $f$  is  $\mu$ -concave. Hence (9) holds for all  $k$  and to prove  $\hat{f}_*^\sigma = \hat{f}_*^\mu = f$  it is sufficient to show that  $\hat{f}_*^\sigma = f$ .

Let  $x_0$  be an arbitrary point in  $\mathcal{A}$ . Then  $x_0 = \sum_{i=1}^{+\infty} \pi_i y_i$ , where  $\{\pi_i\} \in \mathfrak{P}_{+\infty}$  and  $\{y_i\} \in \mathcal{A}_1$ . Let  $x_n = (\lambda_n)^{-1} \sum_{i=1}^n \pi_i y_i$  and  $y_n = (1 - \lambda_n)^{-1} \sum_{i>n} \pi_i y_i$ , where  $\lambda_n = \sum_{i=1}^n \pi_i$ . The sequence  $\{x_n\}$  belongs to the set  $\mathcal{A}_*$  and converges to the point  $x_0$ .

For each  $n$  we have  $x_0 = \lambda_n x_n + (1 - \lambda_n) y_n$  and hence  $\hat{f}_*^\sigma(x_0) \geq \lambda_n \hat{f}_*^\sigma(x_n) = \lambda_n f(x_n)$  by concavity and nonnegativity of the function  $\hat{f}_*^\sigma$ . This implies  $\limsup_{n \rightarrow +\infty} f(x_n) \leq \hat{f}_*^\sigma(x_0)$ . By lower semicontinuity of the function  $f$  we have  $f(x_0) \leq \hat{f}_*^\sigma(x_0)$  and hence  $f(x_0) = \hat{f}_*^\sigma(x_0)$ .  $\square$

Lemma 5B, Proposition 1, Corollary 1B and Corollary 2 imply the following observation, providing usefulness of the approximating sequences  $\{\hat{f}_k^\mu\}$  and  $\{\hat{f}_k^\sigma\}$  for our purposes.

**Proposition 5.** *If the function  $f$  has continuous restriction to the set  $\mathcal{A}_k$  for each  $k$  then the function  $\hat{f}_k^\mu$  is bounded and upper semicontinuous for each  $k$ . If, in addition, the set  $\mathcal{A}$  is stable and  $\mathcal{A} = \sigma\text{-co}(\mathcal{A}_1)$  then  $\hat{f}_k^\sigma = \hat{f}_k^\mu \in C(\mathcal{A})$  for each  $k$ .*

**Remark 4.** The above constructions can be generalized by considering the family  $\{\mathcal{A}_k\}$  produced by an arbitrary closed subset  $\mathcal{A}_1$  of  $\mathcal{A}$  such that  $\mathcal{A} = \overline{\text{co}}(\mathcal{A}_1)$ . The all results remain valid in this case excepting the second assertion of Proposition 5, in which the requirement of openness of the map  $M^a(\mathcal{A}_1) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$  must be added. This can be shown by applying Proposition 3 instead of Corollary 2.

## 5 Continuity conditions

Let  $f$  be a concave nonnegative function on a convex  $\mu$ -compact set  $\mathcal{A}$ . In this section we consider conditions of continuity of this function on subsets of  $\mathcal{A}$  assuming that there exists a closed subset  $\mathcal{A}_1 \subset \mathcal{A}$  such that  $\mathcal{A} = \overline{\text{co}}(\mathcal{A}_1)$  and

$$f|_{\mathcal{A}_k} \text{ is continuous for each natural } k, \tag{13}$$

where  $\mathcal{A}_k$  is the subset of  $\mathcal{A}$  defined by (4). This assumption with  $\mathcal{A}_1 = \text{cl}(\text{extr}\mathcal{A})$  has a physical motivation (see Section 6). Sometimes it can be reduced to continuity and boundedness of  $f|_{\mathcal{A}_1}$  (see the proof of Lemma 6 below).

### 5.1 The case $\mathcal{A} = \sigma\text{-co}(\text{cl}(\text{extr}\mathcal{A}))$

The results of the previous sections imply the following continuity condition.

**Theorem 1.** *Let  $\mathcal{A}$  be a convex  $\mu$ -compact set such that  $\mathcal{A} = \sigma\text{-co}(\text{cl}(\text{extr}\mathcal{A}))$ . Let  $f$  be a concave nonnegative function on the set  $\mathcal{A}$  such that assumption (13) holds with  $\mathcal{A}_1 = \text{cl}(\text{extr}\mathcal{A})$ . Assume that one of the following conditions is valid:*

- a) *the set  $\mathcal{A}$  is stable,*
- b) *the function  $f$  is lower semicontinuous.*

Then the function  $f$  is continuous on a subset  $\mathcal{B} \subseteq \mathcal{A}$  if

$$\lim_{k \rightarrow +\infty} \sup_{x \in \mathcal{B}} \Delta_k^\sigma(x|f) = 0, \quad \text{where} \quad \Delta_k^\sigma(x|f) = \inf_{\{\pi_i, x_i\} \in M_x^a(\mathcal{A}_k)} \left[ f(x) - \sum_i \pi_i f(x_i) \right]. \quad (14)$$

If the both above conditions a) and b) are valid then (14) is a necessary and sufficient condition of continuity of the function  $f$  on a compact subset  $\mathcal{B} \subset \mathcal{A}$ .

**Remark 5.** Since  $\Delta_k^\sigma(x|f) = f - \hat{f}_k^\sigma$  and  $\hat{f}_k^\sigma \leq \hat{f}_k^\mu$ , where  $\hat{f}_k^\mu$  and  $\hat{f}_k^\sigma$  are functions defined by (8) and (11), condition (14) means uniform convergence of the sequences  $\{\hat{f}_k^\sigma\}$  and  $\{\hat{f}_k^\mu\}$  to the function  $f$  on the subset  $\mathcal{B}$ .

**Remark 6.** Applications of the above continuity condition are based on possibility to find for a given concave function  $f$  a suitable upper bound for the value in the square brackets in (14) (see Example 1 below and Section 6).

**Proof.** If the set  $\mathcal{A}$  is stable then  $\hat{f}_k^\mu = \hat{f}_k^\sigma \in C(\mathcal{A})$  for all  $k$  by Proposition 5. By Remark 5 condition (14) implies continuity of the function  $f$  on the subset  $\mathcal{B}$ .

If the function  $f$  is lower semicontinuous then continuity of the function  $f$  on the subset  $\mathcal{B}$  can be verified by showing its upper semicontinuity and boundedness on this set. By Remark 5 the last property follows from condition (14) since by Proposition 5 the sequence  $\{\hat{f}_k^\mu\}$  consists of upper semicontinuous bounded functions.

By Propositions 4 and 5 the last assertion of the theorem follows from Dini's lemma and Remark 5.  $\square$

**Example 1.** The Shannon entropy is a concave lower semicontinuous function on the set  $\mathfrak{P}_{+\infty} = \left\{ x = \{x^j\}_{j=1}^{+\infty} \in \ell_1 \mid x^j \geq 0, \forall j, \sum_{j=1}^{+\infty} x^j = 1 \right\}$  of all countable probability distributions defined as follows

$$S(\{x^j\}_{j=1}^{+\infty}) = - \sum_{j=1}^{+\infty} x^j \ln x^j.$$

This function is nonnegative and takes the value  $+\infty$  on a dense subset of  $\mathfrak{P}_{+\infty}$ .

As mentioned in Section 2 the convex set  $\mathfrak{P}_{+\infty}$  is stable and  $\mu$ -compact. The set  $\text{extr } \mathfrak{P}_{+\infty}$  consists of "degenerate" distributions having "1" at some position and "0" on other places. It is clear that  $\mathfrak{P}_{+\infty} = \sigma\text{-co}(\text{extr } \mathfrak{P}_{+\infty})$  and that the function  $x \mapsto S(x)$  has continuous restriction to the set

$$(\mathfrak{P}_{+\infty})_k = \left\{ \sum_{i=1}^k \pi_i x_i \mid \{\pi_i\} \in \mathfrak{P}_k, \{x_i\} \subset \text{extr } \mathfrak{P}_{+\infty} \right\}$$

for each  $k \in \mathbb{N}$ . If  $f = S$  then the value in the square brackets in (14) can be expressed as follows

$$S(x) - \sum_i \pi_i S(x_i) = \sum_i \pi_i S(x_i \| x),$$

where  $S(\cdot\|\cdot)$  is the relative entropy (Kullback-Leibler distance [18]) defined for arbitrary distributions  $x = \{x^j\}_{j=1}^{+\infty}$  and  $y = \{y^j\}_{j=1}^{+\infty}$  in  $\mathfrak{P}_{+\infty}$  by the formula

$$S(x\|y) = \begin{cases} \sum_{i=1}^{+\infty} x^i \ln(x^i/y^i), & \{y^j = 0\} \Rightarrow \{x^j = 0\} \\ +\infty, & \text{otherwise} \end{cases}.$$

Thus Theorem 1 implies the following continuity condition for the Shannon entropy.

*The function  $x \mapsto S(x)$  is continuous on a compact subset  $\mathfrak{P} \subseteq \mathfrak{P}_{+\infty}$  if and only if*

$$\lim_{k \rightarrow +\infty} \sup_{x \in \mathfrak{P}} \Delta_k^\sigma(x|S) = 0, \quad \text{where} \quad \Delta_k^\sigma(x|S) = \inf_{\{\pi_i, x_i\} \in M_x^\sigma((\mathfrak{P}_{+\infty})_k)} \sum_i \pi_i S(x_i\|x). \quad (15)$$

This condition can be applied directly by using well studied properties of the relative entropy. For example, by joint convexity and lower semicontinuity of the relative entropy validity of (15) for convex subsets  $\mathfrak{P}'$  and  $\mathfrak{P}''$  of  $\mathfrak{P}_{+\infty}$  implies validity of (15) for their convex closure  $\overline{\text{co}}(\mathfrak{P}' \cup \mathfrak{P}'')$ . Hence, we can conclude that *continuity of the Shannon entropy on convex closed subsets  $\mathfrak{P}'$  and  $\mathfrak{P}''$  implies its continuity on their convex closure  $\overline{\text{co}}(\mathfrak{P}' \cup \mathfrak{P}'')$ .*<sup>7</sup>

The above continuity condition can be also applied by using the estimation

$$\Delta_k^\sigma(x|S) \leq S(k(x)), \quad k \in \mathbb{N}, \quad (16)$$

where  $k(x)$  is a distribution obtained by  $k$ -order coarse-graining from the distribution  $x$ , that is  $(k(x))^j = x^{(j-1)k+1} + \dots + x^{jk}$  for all  $j = 1, 2, \dots$ . This estimation is proved by using the decomposition  $x = \sum_{i=1}^{+\infty} \lambda_i^k p_i^k(x)$ , where  $\lambda_i^k = (k(x))^i$  and  $p_i^k(x)$  is a distribution such that  $(p_i^k(x))^j = (\lambda_i^k)^{-1} x^j$  for  $j = (i-1)k+1, ik$  and  $(p_i^k(x))^j = 0$  for others  $j$ , since it is easy to verify that  $\sum_{i=1}^{+\infty} \lambda_i^k S(p_i^k(x)\|x) = \sum_{i=1}^{+\infty} \lambda_i^k (-\ln \lambda_i^k) = S(k(x))$ .

The above continuity condition and estimation (16) imply the following assertion.

*Let  $x_0$  be a distribution in  $\mathfrak{P}_{+\infty}$  with finite Shannon entropy, then the Shannon entropy is continuous on the set*

$$\{x \in \mathfrak{P}_{+\infty} \mid x \prec x_0\}, \quad (17)$$

*where  $x \prec y$  means that the distribution  $y = \{y^j\}_{j=1}^{+\infty}$  is more chaotic than the distribution  $x = \{x^j\}_{j=1}^{+\infty}$  in the Uhlmann sense [2, 32], that is  $\sum_{j=1}^n x^j \geq \sum_{j=1}^n y^j$  for each natural  $n$  provided the sequences  $\{x^j\}_{j=1}^{+\infty}$  and  $\{y^j\}_{j=1}^{+\infty}$  are arranged in nonincreasing order.*<sup>8</sup>

Indeed, assuming that the elements of  $x$  and  $x_0$  are arranged in nonincreasing order we have  $x \prec x_0 \Rightarrow k(x) \prec k(x_0) \Rightarrow S(k(x)) \leq S(k(x_0))$  by Shur concavity of the Shannon entropy [32]. Hence validity of (15) for set (17) follows from (16) and the easily verified implication  $S(x_0) < +\infty \Rightarrow \lim_{k \rightarrow +\infty} S(k(x_0)) = 0$ .

<sup>7</sup>It is possible to show that continuity of the Shannon entropy on a convex subset of  $\mathfrak{P}_{+\infty}$  implies relative compactness of this subset.

<sup>8</sup>The order " $\prec$ " is converse to the majorization order used in linear algebra [6].

## 5.2 Possible generalizations

Note first that Theorem 1 can be generalized by replacing the family  $\{\mathcal{A}_k\}$  produced by the set  $\mathcal{A}_1 = \text{cl}(\text{extr}(\mathcal{A}))$  by a family  $\{\mathcal{A}_k\}$  produced by an arbitrary closed subset  $\mathcal{A}_1$  of  $\mathcal{A}$  such that  $\mathcal{A} = \sigma\text{-co}(\mathcal{A}_1)$ . By Remark 4 the only necessary modification of Theorem 1 under this replacement consists in the additional requirement of openness of the map  $M^a(\mathcal{A}_1) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$  in condition a).

Without the assumption  $\mathcal{A} = \sigma\text{-co}(\mathcal{A}_1)$  the following continuity condition can be proved.

**Theorem 2.** *Let  $\mathcal{A}_1$  be a closed subset of a convex  $\mu$ -compact set  $\mathcal{A}$  such that  $\mathcal{A} = \overline{\text{co}}(\mathcal{A}_1)$  (in particular,  $\mathcal{A}_1 = \text{cl}(\text{extr}(\mathcal{A}))$ ). Let  $f$  be a concave lower semicontinuous nonnegative function on the set  $\mathcal{A}$  satisfying assumption (13). Then the function  $f$  is continuous on a subset  $\mathcal{B} \subseteq \mathcal{A}$  if*

$$\lim_{k \rightarrow +\infty} \sup_{x \in \mathcal{B}} \Delta_k^\mu(x|f) = 0, \quad \text{where} \quad \Delta_k^\mu(x|f) = \inf_{\mu \in M_x^a(\mathcal{A}_k)} \left[ f(x) - \int_{\mathcal{A}_k} f(y) \mu(dy) \right]. \quad (18)$$

Condition (18) can be replaced by the following one

$$\lim_{k \rightarrow +\infty} \sup_{x \in \mathcal{B}_0} \Delta_k^\sigma(x|f) = 0, \quad \text{where} \quad \Delta_k^\sigma(x|f) = \inf_{\{\pi_i, x_i\} \in M_x^\sigma(\mathcal{A}_k)} \left[ f(x) - \sum_i \pi_i f(x_i) \right] \quad (19)$$

and  $\mathcal{B}_0$  is an arbitrary subset of  $\sigma\text{-co}(\mathcal{A}_1)$  such that  $\mathcal{B} \subseteq \text{cl}(\mathcal{B}_0)$ .

**Proof.** By Proposition 5 the function  $\hat{f}_k^\mu$  defined by formula (8) is upper semicontinuous and bounded for each  $k$ . Since  $\Delta_k^\mu(x|f) = f - \hat{f}_k^\mu$ , condition (18) means uniform convergence of the sequence  $\{\hat{f}_k^\mu\}$  to the function  $f$  on the subset  $\mathcal{B}$ , which implies upper semicontinuity and boundedness of the lower semicontinuous function  $f$  on the subset  $\mathcal{B}$ .

Let  $\mathcal{B}_0$  be a subset of  $\sigma\text{-co}(\mathcal{A}_1)$  such that  $\mathcal{B} \subseteq \text{cl}(\mathcal{B}_0)$ . On this subset the function  $\hat{f}_k^\sigma$  is well defined by formula (11) for each  $k$ . Since  $\Delta_k^\sigma(x|f) = f - \hat{f}_k^\sigma$  and  $\hat{f}_k^\sigma \leq \hat{f}_k^\mu$ , condition (19) guarantees uniform convergence of the sequence  $\{\hat{f}_k^\sigma\}$  to the function  $f$  on the subset  $\mathcal{B}_0$ , which implies uniform convergence of the sequence  $\{\hat{f}_k^\mu\}$  to the function  $f$  on the subset  $\text{cl}(\mathcal{B}_0)$ , since the function  $\Delta_k^\mu(x|f) = f - \hat{f}_k^\mu$  is lower semicontinuous (as a difference between lower semicontinuous and bounded upper semicontinuous functions).  $\square$

**Remark 7.** If the set  $\mathcal{A}$  is stable,  $\mathcal{A}_1 = \text{extr}(\mathcal{A})$  and positive answers on the above Questions 1 and 2 (stated respectively in Sections 3 and 4) hold then (18) is a necessary and sufficient condition of continuity of the function  $f$  on a compact subset  $\mathcal{B} \subset \mathcal{A}$ . Necessity of condition (18) in this case can be shown by using Corollary 1A and Dini's lemma.

Theorem 2 can be applied to analysis of concave functions on the stable convex  $\mu$ -compact set of probability measures on a complete separable metric space having continuous restrictions to the subset of measures supported by  $\leq k$  atoms for all  $k$ .

## 6 Applications in quantum physics

The notion of a quantum state plays a central role in the statistical structure of quantum theory [14]. In this section we consider applications of the continuity conditions obtained in the previous section to analysis of local continuity of several entropic characteristics – the particular concave functions on the convex set of all quantum states.

Let  $\mathcal{H}$  be a separable Hilbert space,  $\mathfrak{T}(\mathcal{H})$  – the Banach space of all trace-class operators in  $\mathcal{H}$  with the trace norm, containing the cone  $\mathfrak{T}_+(\mathcal{H})$  of all positive trace-class operators.

The closed convex set

$$\mathfrak{S}(\mathcal{H}) = \{A \in \mathfrak{T}_+(\mathcal{H}) \mid \text{Tr}A = 1\}$$

is a complete separable metric space with the metric defined by the trace norm. Operators in  $\mathfrak{S}(\mathcal{H})$  are denoted  $\rho, \sigma, \omega, \dots$  and called density operators or *quantum states* since each density operator corresponds to a normal state on the algebra of all bounded operators [8].

It is essential that the convex set  $\mathfrak{S}(\mathcal{H})$  is stable and  $\mu$ -compact [28] (the set  $\mathfrak{S}(\mathcal{H})$  is compact if and only if  $\dim \mathcal{H} < +\infty$ ). The set  $\text{extr}\mathfrak{S}(\mathcal{H})$  of its extreme points consists of one dimensional projectors – *pure* states. A pure state corresponding to a unit vector  $|\varphi\rangle \in \mathcal{H}$  will be denoted  $|\varphi\rangle\langle\varphi|$ . By the spectral theorem an arbitrary state  $\rho$  can be represented as follows  $\rho = \sum_i \lambda_i |\varphi_i\rangle\langle\varphi_i|$ , where  $\{|\varphi_i\rangle\}$  is the orthonormal basis of eigenvectors of the operator  $\rho$  and  $\{\lambda_i\}$  is the corresponding sequence of eigenvalues. Hence  $\mathfrak{S}(\mathcal{H}) = \sigma\text{-co}(\text{extr}\mathfrak{S}(\mathcal{H}))$ .

Rapid development of quantum information theory leads to discovery of a whole number of important entropic and informational characteristics of quantum systems, see e.g. [14, 21]. Many of them can be considered as functions on the set of quantum states. In the finite dimensional case ( $\dim \mathcal{H} < +\infty$ ) these functions are generally bounded and continuous on the whole set of quantum states, but in infinite dimensions their analytical properties are not so good. For example, the von Neumann entropy is a continuous bounded function on the set of quantum states of finite dimensional quantum system, but it is discontinuous and takes the value  $+\infty$  on a dense subset of the set of all infinite dimensional quantum states.<sup>9</sup>

Discontinuity and unboundedness of entropic characteristics lead to technical problems in analysis of infinite dimensional quantum systems. Moreover, they produce a number of "nonphysical" effects such as infinite values of different capacities of a quantum channel and their discontinuity as functions of a channel [19, 26]. But these difficulties can be partially overcome by using local continuity conditions for entropic characteristics [20, 26, 31]. For example, continuity of the von Neumann entropy on the set of states of the system of quantum oscillators with bounded mean energy provides many results concerning different characteristics of this system (see [13] and references therein). Thus, study of local continuity properties of entropic characteristics of quantum states is important for rigorous analysis of infinite dimensional quantum systems.

Since  $\mathfrak{S}(\mathcal{H}) = \sigma\text{-co}(\text{extr}\mathfrak{S}(\mathcal{H}))$  is a convex stable  $\mu$ -compact set, we can apply the results of the previous sections to study concave nonnegative functions on the set  $\mathfrak{S}(\mathcal{H})$  having

---

<sup>9</sup>Moreover, the set of states with finite von Neumann entropy is a first category subset of the set of all quantum states [31].

restrictions to the set

$$\mathfrak{S}_k(\mathcal{H}) = \left\{ \sum_{i=1}^k \pi_i \rho_i \mid \{\pi_i\} \in \mathfrak{P}_k, \{\rho_i\} \subset \text{extr} \mathfrak{S}(\mathcal{H}) \right\} \quad (20)$$

with appropriate analytical properties for all  $k$ . Note that  $\mathfrak{S}_k(\mathcal{H})$  is the set of all quantum states having rank  $\leq k$  (as operators in  $\mathcal{H}$ ), it can be considered as an union of all unitary translations of the set  $\mathfrak{S}(\mathcal{H}_k)$ , where  $\mathcal{H}_k$  is a particular  $k$ -dimensional subspace of  $\mathcal{H}$ .

Let  $f$  be a concave nonnegative function on the set  $\mathfrak{S}(\mathcal{H})$ . For given natural  $k$  consider the concave functions

$$\hat{f}_k^\mu(\rho) = \sup_{\mu \in M_\rho(\mathfrak{S}_k(\mathcal{H}))} \int_{\mathfrak{S}_k(\mathcal{H})} f(\sigma) \mu(d\sigma) \quad \text{and} \quad \hat{f}_k^\sigma(\rho) = \sup_{\{\pi_i, \rho_i\} \in M_\rho^a(\mathfrak{S}_k(\mathcal{H}))} \sum_i \pi_i f(\rho_i)$$

on the set  $\mathfrak{S}(\mathcal{H})$  (assuming that  $f$  has universally measurable restriction to the set  $\mathfrak{S}_k(\mathcal{H})$ ).

It is clear that  $\hat{f}_k^\sigma \leq \hat{f}_k^\mu$ . Since the function  $f$  is  $\sigma$ -concave by Lemma 1A we have  $\hat{f}_k^\sigma \leq f$  and  $\hat{f}_k^\sigma|_{\mathfrak{S}_k(\mathcal{H})} = f|_{\mathfrak{S}_k(\mathcal{H})}$ . If the function  $f$  is  $\mu$ -concave (see conditions in Lemma 1B) then  $\hat{f}_k^\mu \leq f$  and  $\hat{f}_k^\mu|_{\mathfrak{S}_k(\mathcal{H})} = f|_{\mathfrak{S}_k(\mathcal{H})}$ .

The results of Sections 3 and 4 imply the following observations.

**Proposition 6.** *Let  $f$  be a concave nonnegative function on the set  $\mathfrak{S}(\mathcal{H})$ , taking finite value at least at one state.*

- A) *If  $f|_{\mathfrak{S}_k(\mathcal{H})}$  is upper semicontinuous for each  $k$  then the function  $\hat{f}_k^\mu$  is upper semicontinuous and bounded on the set  $\mathfrak{S}(\mathcal{H})$  for each  $k$ .*
- B) *If  $f|_{\mathfrak{S}_k(\mathcal{H})}$  is lower semicontinuous for each  $k$  then  $\hat{f}_k^\mu = \hat{f}_k^\sigma$  and this function is lower semicontinuous on the set  $\mathfrak{S}(\mathcal{H})$  for each  $k$ .*
- C) *If  $f|_{\mathfrak{S}_k(\mathcal{H})}$  is continuous for each  $k$  then  $\hat{f}_k^\mu = \hat{f}_k^\sigma \in C(\mathfrak{S}(\mathcal{H}))$  for each  $k$ .*

*If the function  $f$  is lower semicontinuous on the set  $\mathfrak{S}(\mathcal{H})$  then the nondecreasing sequence  $\{\hat{f}_k^\mu = \hat{f}_k^\sigma\}$  pointwise converges to the function  $f$ .*

By Proposition 6 an arbitrary concave lower semicontinuous nonnegative function  $f$  on the set  $\mathfrak{S}(\mathcal{H})$  having continuous restriction to the set  $\mathfrak{S}_k(\mathcal{H})$  for each  $k$  can be approximated by the increasing sequence of concave continuous nonnegative bounded functions  $f_k \doteq \hat{f}_k^\mu = \hat{f}_k^\sigma$  such that  $f_k|_{\mathfrak{S}_k(\mathcal{H})} = f|_{\mathfrak{S}_k(\mathcal{H})}$  for each  $k$ . Advantages of this approximation and its possible applications are considered in [27, Section 4] and in [28, Section 6.2].

Theorem 1 implies the following continuity condition (extending the results of [27]).

**Proposition 7.** *Let  $f$  be a concave nonnegative function on the set  $\mathfrak{S}(\mathcal{H})$  having continuous restriction to the set  $\mathfrak{S}_k(\mathcal{H})$  defined by (20) for each  $k$ . Then the function  $f$  is continuous on a subset  $\mathfrak{S} \subseteq \mathfrak{S}(\mathcal{H})$  if*

$$\lim_{k \rightarrow +\infty} \sup_{\rho \in \mathfrak{S}} \Delta_k^\sigma(\rho|f) = 0, \quad \text{where} \quad \Delta_k^\sigma(\rho|f) = \inf_{\{\pi_i, \rho_i\} \in M_\rho^a(\mathfrak{S}_k(\mathcal{H}))} \left[ f(\rho) - \sum_i \pi_i f(\rho_i) \right]. \quad (21)$$

If the function  $f$  is lower semicontinuous then (21) is a necessary and sufficient condition of continuity of the function  $f$  on a compact subset  $\mathfrak{S} \subset \mathfrak{S}(\mathcal{H})$ .

The conditions of Proposition 7 are valid for the following well known characteristics of quantum states – concave lower semicontinuous nonnegative functions on the set  $\mathfrak{S}(\mathcal{H})$ :

- the quantum Renyi entropy  $R_p(\rho) = \ln \text{Tr} \rho^p / (1 - p)$  of order  $p \in (0, 1]$  (the case  $p = 1$  corresponds to the von Neumann entropy  $H(\rho) = -\text{Tr} \rho \ln \rho$ );
- the quantum mutual information  $I(\rho, \Phi)$  of a quantum channel  $\Phi$  (defined in Section 6.2);
- the output quantum Renyi entropy  $R_p(\Phi(\rho))$  of order  $p \in (0, 1]$  (in particular, the output von Neumann entropy  $H(\Phi(\rho))$ ) of a quantum channel  $\Phi$  satisfying the particular condition (see Section 6.3).
- the  $\chi$ -function (the constrained Holevo capacity)  $\chi_\Phi(\rho)$  a quantum channel  $\Phi$  satisfying the particular condition (see Section 6.3).

Below we consider applications of Proposition 7 to the above functions, reducing attention to the von Neumann entropy – the most important version of the quantum Renyi entropy.

## 6.1 The von Neumann entropy

Continuity conditions for the von Neumann entropy on subsets of  $\mathfrak{T}_+(\mathcal{H})$  based on the above approximation technic are presented in [27]. Here we consider the case of the von Neumann entropy for completeness, reducing attention to subsets of  $\mathfrak{S}(\mathcal{H})$ .

The von Neumann entropy  $H(\rho) = -\text{Tr} \rho \ln \rho$  is a concave lower semicontinuous unitary invariant function on the set  $\mathfrak{S}(\mathcal{H})$  taking values in  $[0, +\infty]$ . It obviously has continuous restriction to the set  $\mathfrak{S}_k(\mathcal{H})$  for each  $k$ . If  $f = H$  then the value in the square brackets in (21) can be expressed as follows

$$H(\rho) - \sum_i \pi_i H(\rho_i) = \sum_i \pi_i H(\rho_i \| \rho), \quad (22)$$

where  $H(\cdot \| \cdot)$  is the quantum relative entropy defined for arbitrary states  $\rho$  and  $\sigma$  in  $\mathfrak{S}(\mathcal{H})$  by the formula

$$H(\rho \| \sigma) = \begin{cases} \sum_{i=1}^{+\infty} \langle \varphi_i | \rho \ln \rho - \rho \ln \sigma | \varphi_i \rangle, & \text{supp } \rho \subseteq \text{supp } \sigma \\ +\infty, & \text{supp } \rho \not\subseteq \text{supp } \sigma \end{cases},$$

in which  $\{|\varphi_i\rangle\}_{i=1}^{+\infty}$  is the orthonormal basis of eigenvectors of the operator  $\rho$  (or  $\sigma$ ) and  $\text{supp } \rho = \mathcal{H} \ominus \ker \rho$  [20, 31]. Thus we obtain from Proposition 7 the following continuity condition for the von Neumann entropy.



**Corollary 3.** *The function  $\rho \mapsto H(\rho)$  is continuous on a compact subset  $\mathfrak{S} \subset \mathfrak{S}(\mathcal{H})$  if and only if*

$$\lim_{k \rightarrow +\infty} \sup_{\rho \in \mathfrak{S}} \Delta_k^\sigma(\rho|H) = 0, \quad \text{where} \quad \Delta_k^\sigma(\rho|H) = \inf_{\{\pi_i, \rho_i\} \in M_\rho^\sigma(\mathfrak{S}_k(\mathcal{H}))} \sum_i \pi_i H(\rho_i|\rho). \quad (23)$$

In [27] the property of an arbitrary subset  $\mathfrak{S} \subseteq \mathfrak{S}(\mathcal{H})$  expressed by (23) is called the uniform approximation property (briefly, the UA-property) and is studied in detail (in the extended context of the positive cone  $\mathfrak{T}_+(\mathcal{H})$  instead of the set  $\mathfrak{S}(\mathcal{H})$ ). By Corollary 3 the UA-property of an arbitrary subset  $\mathfrak{S}$  is a sufficient condition of continuity of the von Neumann entropy on this subset and this condition is necessary if the subset  $\mathfrak{S}$  is compact.

Usefulness of the UA-property as a continuity condition is based on possibility to analyze it by applying well studied properties of the quantum relative entropy. This makes it possible to find a class of different set-operations preserving the UA-property ([27, Proposition 4]). For example,

- by joint convexity and lower semicontinuity of the quantum relative entropy the UA-property of convex subsets  $\mathfrak{S}_1$  and  $\mathfrak{S}_2$  of  $\mathfrak{S}(\mathcal{H})$  implies the UA-property of their convex closure  $\overline{\text{co}}(\mathfrak{S}_1 \cup \mathfrak{S}_2)$ ;
- by monotonicity of the quantum relative entropy the UA-property a subset  $\mathfrak{S}$  of  $\mathfrak{S}(\mathcal{H})$  implies the UA-property of the set  $\{\Phi(\rho) \mid \Phi \in \mathfrak{F}_n, \rho \in \mathfrak{S}\}$ , where  $\mathfrak{F}_n$  is the set of all quantum channels having the Kraus representation consisting of  $\leq n$  summands.<sup>10</sup>

By using the first above-stated observation it is easy to show that continuity of the von Neumann entropy on convex closed subsets  $\mathfrak{S}_1$  and  $\mathfrak{S}_2$  of  $\mathfrak{S}(\mathcal{H})$  implies its continuity on their convex closure  $\overline{\text{co}}(\mathfrak{S}_1 \cup \mathfrak{S}_2)$  (Corollary 7 in [27]), while the second one implies the result concerning continuity of the von Neumann entropy of posteriority states in quantum measurements (Example 3 in [27]).

The continuity condition based on the UA-property gives the universal method of proving continuity of the von Neumann entropy. Various applications of this method are considered in [27, Section 5.2]. The "only if" part of Corollary 3 makes it possible to prove that continuity of the von Neumann entropy on some set of states implies continuity of other important entropic characteristics on this set (see the proofs of Corollaries 4 and 5 below).

## 6.2 The quantum mutual information

Let  $\mathcal{H}$  and  $\mathcal{H}'$  be two separable Hilbert spaces. A completely positive trace-preserving linear map  $\Phi : \mathfrak{T}(\mathcal{H}) \rightarrow \mathfrak{T}(\mathcal{H}')$  is called *quantum channel* [14, 21]. By the Stinespring dilation theorem there exist a separable Hilbert space  $\mathcal{H}''$  and an isometry  $V : \mathcal{H} \rightarrow \mathcal{H}' \otimes \mathcal{H}''$  such that

$$\Phi(A) = \text{Tr}_{\mathcal{H}''} V A V^*, \quad \forall A \in \mathfrak{T}(\mathcal{H}). \quad (24)$$

---

<sup>10</sup>The notions of a quantum channel and of its Kraus representation are described in the next subsection.

The quantum channel

$$\mathfrak{T}(\mathcal{H}) \ni A \mapsto \tilde{\Phi}(A) = \text{Tr}_{\mathcal{H}'} V A V^* \in \mathfrak{T}(\mathcal{H}'') \quad (25)$$

is called *complementary* to the channel  $\Phi$ , it is uniquely defined up to unitary equivalence [15]. By using representation (24) it is easy to obtain the Kraus representation

$$\Phi(A) = \sum_{j=1}^{+\infty} V_j A V_j^*, \quad \forall A \in \mathfrak{T}(\mathcal{H}), \quad (26)$$

where  $\{V_j\}_{j=1}^{+\infty}$  is a set of bounded linear operators from  $\mathcal{H}$  to  $\mathcal{H}'$  such that  $\sum_{j=1}^{+\infty} V_j^* V_j = I_{\mathcal{H}}$ . Via the set  $\{V_j\}_{j=1}^{+\infty}$  of Kraus operators of the channel  $\Phi$  its complementary channel can be expressed as follows

$$\tilde{\Phi}(A) = \sum_{i,j=1}^{+\infty} \text{Tr} [V_i A V_j^*] |\varphi_i\rangle\langle\varphi_j|, \quad A \in \mathfrak{T}(\mathcal{H}), \quad (27)$$

where  $\{|\varphi_i\rangle\}_{i=1}^{+\infty}$  is a particular orthonormal basis in the space  $\mathcal{H}''$  [15].

In finite dimensions ( $\dim \mathcal{H}, \dim \mathcal{H}' < +\infty$ ) the quantum mutual information of the channel  $\Phi$  at a state  $\rho \in \mathfrak{S}(\mathcal{H})$  is defined as follows (cf.[1])

$$I(\rho, \Phi) = H(\rho) + H(\Phi(\rho)) - H(\tilde{\Phi}(\rho)). \quad (28)$$

This is an important characteristic of a quantum channel related to the entanglement-assisted classical capacity of this channel [5].

In infinite dimensions the above definition may contain the uncertainty " $\infty - \infty$ ", but it can be modified to avoid this problem as follows

$$I(\rho, \Phi) = H(\Phi \otimes \text{Id}_{\mathcal{K}}(|\varphi_{\rho}\rangle\langle\varphi_{\rho}|) \| \Phi(\rho) \otimes \rho), \quad (29)$$

where  $\mathcal{K} \cong \mathcal{H}$ ,  $\text{Id}_{\mathcal{K}}$  is the identity transformation of  $\mathfrak{T}(\mathcal{H})$  and  $|\varphi_{\rho}\rangle$  is a unit vector in  $\mathcal{H} \otimes \mathcal{K}$  such that  $\text{Tr}_{\mathcal{K}} |\varphi_{\rho}\rangle\langle\varphi_{\rho}| = \text{Tr}_{\mathcal{H}} |\varphi_{\rho}\rangle\langle\varphi_{\rho}| = \rho$ . In [16] it is shown that for an arbitrary quantum channel  $\Phi$  the nonnegative function  $\rho \mapsto I(\rho, \Phi)$  defined by (29) is concave and lower semicontinuous on the set  $\mathfrak{S}(\mathcal{H})$  (Proposition 1) and that this function is continuous on each subset of  $\mathfrak{S}(\mathcal{H})$  on which the von Neumann entropy is continuous, in particular, it is continuous on the set  $\mathfrak{S}_k(\mathcal{H})$  for each  $k$  (Proposition 4). Hence for an arbitrary quantum channel  $\Phi$  the conditions of Proposition 7 are valid for the function  $\rho \mapsto I(\rho, \Phi)$ .

By using identity (22), formula (28) and a simple approximation it is possible to show that

$$\Delta_k^{\sigma}(\rho | I_{\Phi}) = \inf_{\{\pi_i, \rho_i\} \in M_{\rho}^{\sigma}(\mathfrak{S}_k(\mathcal{H}))} \sum_i \pi_i \left[ H(\rho_i \| \rho) + H(\Phi(\rho_i) \| \Phi(\rho)) - H(\tilde{\Phi}(\rho_i) \| \tilde{\Phi}(\rho)) \right], \quad (30)$$

where  $I_{\Phi}(\cdot) \doteq I(\cdot, \Phi)$ , for any state  $\rho$  in  $\mathfrak{S}(\mathcal{H})$  with finite entropy. The expression in the right side of (30) is well defined, since  $\sum_i \pi_i H(\tilde{\Phi}(\rho_i) \| \tilde{\Phi}(\rho)) \leq \sum_i \pi_i H(\rho_i \| \rho) \leq H(\rho)$  by

monotonicity of the quantum relative entropy and identity (22).

Proposition 7 and Corollary 3 imply the following continuity condition for the quantum mutual information, strengthening Proposition 4 in [16].

**Corollary 4.** *Let  $\Phi$  be an arbitrary quantum channel and  $\mathfrak{S}$  be a compact subset of  $\mathfrak{S}(\mathcal{H})$  on which the von Neumann entropy is finite. The following assertions*

- (i) *the function  $\rho \mapsto H(\rho)$  is continuous on the set  $\mathfrak{S}$ ,*
- (ii)  *$\lim_{k \rightarrow +\infty} \sup_{\rho \in \mathfrak{S}} \Delta_k^\sigma(\rho|I_\Phi) = 0$ , where  $\Delta_k^\sigma(\rho|I_\Phi)$  is defined by (30),*
- (iii) *the function  $\rho \mapsto I(\rho, \Phi)$  is continuous on the set  $\mathfrak{S}$ ,*

*are related by the implications (i)  $\Rightarrow$  (ii)  $\Leftrightarrow$  (iii).*

*If  $\Phi$  is a degradable channel, that is  $\tilde{\Phi} = \Lambda \circ \Phi$  for some quantum channel  $\Lambda$ , then assertions (i) – (iii) are equivalent for an arbitrary compact subset  $\mathfrak{S}$  of  $\mathfrak{S}(\mathcal{H})$ .*

**Proof.** (i)  $\Rightarrow$  (ii) is proved by using Corollary 3, since by monotonicity and nonnegativity of the quantum relative entropy the expression in the square brackets in (30) does not exceed  $2H(\rho_i||\rho)$  and hence (ii) follows from (23). (ii)  $\Leftrightarrow$  (iii) follows from Proposition 7.

If  $\Phi$  is a degradable channel then by using Theorem 1 in [16] and the 1-th chain rule from Proposition 1 in [16] it is easy to show that  $I(\rho, \Phi) < +\infty \Rightarrow H(\rho) < +\infty$ , while by monotonicity of the quantum relative entropy the expression in the square brackets in (30) is not less than  $H(\rho_i||\rho)$ . Thus (23) follows from (ii) in this case.  $\square$

### 6.3 The output von Neumann entropy and the $\chi$ -function of a quantum channel

Let  $\Phi : \mathfrak{T}(\mathcal{H}) \rightarrow \mathfrak{T}(\mathcal{H}')$  be a quantum channel (see Section 6.2). The output von Neumann entropy  $H(\Phi(\cdot))$  is an important characteristic involved, in particular, in expressions for different capacities of this channel (directly or via other characteristics) [14, 21].

The function  $\rho \mapsto H_\Phi(\rho) \doteq H(\Phi(\rho))$  is concave lower semicontinuous and nonnegative on the set  $\mathfrak{S}(\mathcal{H})$ , but in general this function is not continuous on sets of the family  $\{\mathfrak{S}_k(\mathcal{H})\}$ . To apply Proposition 7 to the function  $\rho \mapsto H_\Phi(\rho)$  we need the following lemma.

**Lemma 6.** *If the function  $\rho \mapsto H_\Phi(\rho)$  is continuous and bounded on the set  $\text{extr}\mathfrak{S}(\mathcal{H})$  then this function is continuous on the set  $\mathfrak{S}_k(\mathcal{H})$  defined by (20) for each natural  $k$ .*

**Proof.** Suppose there exists a sequence  $\{\rho_n\} \subset \mathfrak{S}_k(\mathcal{H})$  converging to a state  $\rho_0 \in \mathfrak{S}_k(\mathcal{H})$  such that

$$\lim_{n \rightarrow +\infty} H_\Phi(\rho_n) > H_\Phi(\rho_0). \quad (31)$$

For each  $n$  we have  $\rho_n = \sum_{i=1}^k \lambda_i^n \sigma_i^n$ , where  $\{\sigma_i^n\}_{i=1}^k \subset \text{extr}\mathfrak{S}(\mathcal{H})$  and  $\{\lambda_i^n\}_{i=1}^k \in \underline{\mathfrak{P}}_k$ . By Lemma 4 we may consider that there exists  $\lim_{n \rightarrow +\infty} \lambda_i^n \sigma_i^n = A_i$  for each  $i = 1, k$ , where  $\{A_i\}_{i=1}^k$  is a set positive trace class operators of rank  $\leq 1$  such that  $\rho_0 = \sum_{i=1}^k A_i$ . Continuity

and boundedness of the function  $H_\Phi$  on the set  $\text{extr}\mathfrak{S}(\mathcal{H})$  imply continuity of its natural extension to the cone of positive trace class operators of rank  $\leq 1$  defined as follows

$$H_\Phi(A) = \text{Tr}AH_\Phi\left(\frac{A}{\text{Tr}A}\right) = \text{Tr}\eta(\Phi(A)) - \eta(\text{Tr}A), \quad A \in \mathfrak{T}_+(\mathcal{H}), \quad \eta(x) = -x \ln x.$$

Hence

$$\lim_{n \rightarrow +\infty} H_\Phi(\lambda_i^n \sigma_i^n) = H_\Phi(A_i), \quad i = \overline{1, k}.$$

By using the property of the von Neumann entropy presented after Corollary 4 in [27] we obtain a contradiction to (31).  $\square$

The  $\chi$ -function of a quantum channel  $\Phi : \mathfrak{T}(\mathcal{H}) \rightarrow \mathfrak{T}(\mathcal{H}')$  is a characteristic related to the classical capacity of this channel [14, 21]. It is defined as follows

$$\chi_\Phi(\rho) = \sup_{\{\pi_i, \rho_i\} \in M^a(\mathfrak{S}(\mathcal{H}))} \sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho)), \quad \rho \in \mathfrak{S}(\mathcal{H}).$$

For a given subset  $\mathfrak{S}$  of  $\mathfrak{S}(\mathcal{H})$  the value  $\sup_{\rho \in \mathfrak{S}} \chi_\Phi(\rho)$  coincides with the Holevo capacity of the  $\mathfrak{S}$ -constrained channel  $\Phi$  [26].

The function  $\rho \mapsto \chi_\Phi(\rho)$  is obviously concave and nonnegative on the set  $\mathfrak{S}(\mathcal{H})$ . In [26] it is proved that this function is lower semicontinuous (Proposition 4) and has continuous restriction to any subset of  $\mathfrak{S}(\mathcal{H})$  on which the function  $\rho \mapsto H_\Phi(\rho)$  is continuous (Theorem 1). Hence Lemma 6 shows that the function  $\rho \mapsto \chi_\Phi(\rho)$  has continuous restriction to the set  $\mathfrak{S}_k(\mathcal{H})$  for each  $k$  if the function  $\rho \mapsto H_\Phi(\rho)$  is continuous and bounded on the set  $\text{extr}\mathfrak{S}(\mathcal{H})$ .

Thus Proposition 7 with Lemma 6 and Corollary 3 imply the following observation.

**Corollary 5.** *Let  $\Phi$  be a quantum channel such that the function  $\rho \mapsto H_\Phi(\rho)$  is continuous and bounded on the set  $\text{extr}\mathfrak{S}(\mathcal{H})$ . Let  $\mathfrak{S}$  be a compact subset of  $\mathfrak{S}(\mathcal{H})$ . The following assertions*

(i) *the function  $\rho \mapsto H(\rho)$  is continuous on the set  $\mathfrak{S}$ ,*

(ii)  $\lim_{k \rightarrow +\infty} \sup_{\rho \in \mathfrak{S}} \Delta_k^\sigma(\rho | H_\Phi) = 0$ , *where  $\Delta_k^\sigma(\rho | H_\Phi) = \inf_{\{\pi_i, \rho_i\} \in M_\rho^\sigma(\mathfrak{S}_k(\mathcal{H}))} \sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho))$ ,*

(iii) *the function  $\rho \mapsto H_\Phi(\rho)$  is continuous on the set  $\mathfrak{S}$ ,*

(iv) *the function  $\rho \mapsto \chi_\Phi(\rho)$  is continuous on the set  $\mathfrak{S}$ ,*

*are related by the implications (i)  $\Rightarrow$  (ii)  $\Leftrightarrow$  (iii)  $\Leftrightarrow$  (iv).*

*If the output entropy of the complementary channel  $\tilde{\Phi}$  is continuous on the set  $\mathfrak{S}$  then assertions (i) – (iv) are equivalent.*

By Corollary 5 the above assertions (i) – (iv) are equivalent for arbitrary quantum channel  $\Phi$  having Kraus representation (26) with finite nonzero summands.

**Proof.** (i)  $\Rightarrow$  (ii) follows from Corollary 3 since monotonicity of the quantum relative entropy implies

$$\inf_{\{\pi_i, \rho_i\} \in M_\rho^a(\mathfrak{S}_k(\mathcal{H}))} \sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho)) \leq \inf_{\{\pi_i, \rho_i\} \in M_\rho^a(\mathfrak{S}_k(\mathcal{H}))} \sum_i \pi_i H(\rho_i \| \rho), \quad \forall k.$$

(ii)  $\Leftrightarrow$  (iii) can be shown by applying Proposition 7 to the function  $\rho \mapsto H_\Phi(\rho)$  and by using the identity

$$H_\Phi(\rho) - \sum_i \pi_i H_\Phi(\rho_i) = \sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho)).$$

(iii)  $\Rightarrow$  (iv) follows from Theorem 1 in [26].

(iv)  $\Rightarrow$  (ii) can be shown by applying Proposition 7 to the function  $\rho \mapsto \chi_\Phi(\rho)$  and by using the inequality

$$\chi_\Phi(\rho) - \sum_i \pi_i \chi_\Phi(\rho_i) \geq \sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho)),$$

valid for any  $\{\pi_i, \rho_i\} \in M^a(\mathfrak{S}(\mathcal{H}))$  [26, Proposition 4].

To prove the last assertion of the corollary it suffices to show that continuity of the both functions  $\rho \mapsto H_\Phi(\rho)$  and  $\rho \mapsto H_{\tilde{\Phi}}(\rho)$  on the set  $\mathfrak{S}$  implies continuity of the function  $\rho \mapsto H(\rho)$  on this set. This can be done by using Lemma 7 below and representations (24) and (25).  $\square$

**Lemma 7.** *Let  $\{\omega_n\}$  be a sequence of states in  $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$  converging to a state  $\omega_0$ . If  $\lim_{n \rightarrow +\infty} H(\text{Tr}_{\mathcal{K}} \omega_n) = H(\text{Tr}_{\mathcal{K}} \omega_0) < +\infty$  and  $\lim_{n \rightarrow +\infty} H(\text{Tr}_{\mathcal{H}} \omega_n) = H(\text{Tr}_{\mathcal{H}} \omega_0) < +\infty$  then  $\lim_{n \rightarrow +\infty} H(\omega_n) = H(\omega_0) < +\infty$ .*

**Proof.** Let  $\omega_n^{\mathcal{H}} \doteq \text{Tr}_{\mathcal{K}} \omega_n$  and  $\omega_n^{\mathcal{K}} \doteq \text{Tr}_{\mathcal{H}} \omega_n$  for  $n = 0, 1, 2, \dots$  Since

$$H(\omega_n) = H(\omega_n^{\mathcal{H}}) + H(\omega_n^{\mathcal{K}}) - H(\omega_n \| \omega_n^{\mathcal{H}} \otimes \omega_n^{\mathcal{K}}),$$

we may consider that  $H(\omega_n) < +\infty$  for  $n = 0, 1, 2, \dots$  and by lower semicontinuity of the quantum relative entropy we have

$$\begin{aligned} \limsup_{n \rightarrow +\infty} H(\omega_n) &= \lim_{n \rightarrow +\infty} H(\omega_n^{\mathcal{H}}) + \lim_{n \rightarrow +\infty} H(\omega_n^{\mathcal{K}}) - \liminf_{n \rightarrow +\infty} H(\omega_n \| \omega_n^{\mathcal{H}} \otimes \omega_n^{\mathcal{K}}) \\ &\leq H(\omega_0^{\mathcal{H}}) + H(\omega_0^{\mathcal{K}}) - H(\omega_0 \| \omega_0^{\mathcal{H}} \otimes \omega_0^{\mathcal{K}}) = H(\omega_0). \end{aligned}$$

This and lower semicontinuity of the von Neumann entropy imply  $\lim_{n \rightarrow +\infty} H(\omega_n) = H(\omega_0)$ .  $\square$

Corollary 4 and Lemma 5B imply the following observation.

**Corollary 6.** *Let  $\Phi$  be a quantum channel. The following assertions are equivalent:*

- (i) *the function  $\rho \mapsto H_\Phi(\rho)$  is continuous and bounded on the set  $\text{extr} \mathfrak{S}(\mathcal{H})$ ,*
- (ii) *the function  $\rho \mapsto H_\Phi(\rho)$  is continuous on any subset of  $\mathfrak{S}(\mathcal{H})$  on which the von Neumann entropy is continuous.*

If assertion (ii) in Corollary 6 holds for a quantum channel one can say roughly speaking that this channel *preserves continuity* of the von Neumann entropy. Assertion (i) in Corollary 6 can be considered as a criterion of this property. It implies, in particular, that the class of quantum channels preserving continuity of the von Neumann entropy contains all quantum channels having Kraus representation (26) with finite nonzero summands. The above criterion also shows that this class contains a quantum channel  $\Phi$  if and only if it contains the complementary channel  $\tilde{\Phi}$  (since  $H_\Phi(\rho) = H_{\tilde{\Phi}}(\rho)$  for any  $\rho \in \text{extr}\mathfrak{S}(\mathcal{H})$  [15]).

## 7 Appendix

**The proof of Proposition 1.** The function  $\hat{f}_\mathcal{B}^\mu$  is well defined on the set  $\overline{\text{co}}(\mathcal{B})$  by Lemma 2. Concavity of the function  $\hat{f}_\mathcal{B}^\mu$  follows from its definition and convexity of the set  $M(\mathcal{B})$ . By upper semicontinuity of functional  $\mathbf{f}$  defined by (2) and compactness of the set  $M_x(\mathcal{B})$  for each  $x$  in  $\overline{\text{co}}(\mathcal{B})$  (provided by  $\mu$ -compactness of the set  $\mathcal{A}$ ) the supremum in the definition of the value  $\hat{f}_\mathcal{B}^\mu(x)$  is achieved at a particular measure  $\mu_x$  in  $M_x(\mathcal{B})$ , that is  $\hat{f}_\mathcal{B}^\mu(x) = \mathbf{f}(\mu_x)$ .

Suppose the function  $\hat{f}_\mathcal{B}^\mu$  is not upper semicontinuous. Then there exists a sequence  $\{x_n\} \subset \overline{\text{co}}(\mathcal{B})$  converging to a point  $x_0 \in \overline{\text{co}}(\mathcal{B})$  such that

$$\exists \lim_{n \rightarrow +\infty} \hat{f}_\mathcal{B}^\mu(x_n) > \hat{f}_\mathcal{B}^\mu(x_0). \quad (32)$$

As proved before for each  $n$  there exists a measure  $\mu_n \in M_{x_n}(\mathcal{B})$  such that  $\hat{f}_\mathcal{B}^\mu(x_n) = \mathbf{f}(\mu_n)$ . The  $\mu$ -compactness of the set  $\mathcal{A}$  implies existence of a subsequence  $\{\mu_{n_k}\}$  converging to a particular measure  $\mu_0$  in  $M(\mathcal{B})$ . By continuity of the map  $\mu \mapsto \mathbf{b}(\mu)$  the measure  $\mu_0$  belongs to the set  $M_{x_0}(\mathcal{B})$ . Upper semicontinuity of the functional  $\mathbf{f}$  implies

$$\hat{f}_\mathcal{B}^\mu(x_0) \geq \mathbf{f}(\mu_0) \geq \limsup_{k \rightarrow +\infty} \mathbf{f}(\mu_{n_k}) = \lim_{k \rightarrow +\infty} \hat{f}_\mathcal{B}^\mu(x_{n_k}),$$

contradicting to (32).

Upper semicontinuity of the concave function  $\hat{f}_\mathcal{B}^\mu$  implies its  $\mu$ -concavity by Lemma 1.

**The proof of Proposition 2.** A) Suppose the function  $\hat{f}_\mathcal{B}^\mu$  is not lower semicontinuous. Then there exists a sequence  $\{x_n\} \subset \overline{\text{co}}(\mathcal{B})$  converging to a point  $x_0 \in \overline{\text{co}}(\mathcal{B})$  such that

$$\exists \lim_{n \rightarrow +\infty} \hat{f}_\mathcal{B}^\mu(x_n) < \hat{f}_\mathcal{B}^\mu(x_0). \quad (33)$$

For  $\varepsilon > 0$  let  $\mu_0^\varepsilon$  be a measure in  $M_{x_0}(\mathcal{B})$  such that  $\hat{f}_\mathcal{B}^\mu(x_0) \leq \mathbf{f}(\mu_0^\varepsilon) + \varepsilon$  ( $\mathbf{f}$  is the functional defined by (2)). By openness of the map  $M(\mathcal{B}) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathcal{A}$  there exists a subsequence  $\{x_{n_k}\}$  and a sequence  $\{\mu_k\} \subset M(\mathcal{B})$  converging to the measure  $\mu_0^\varepsilon$  such that  $\mathbf{b}(\mu_k) = x_{n_k}$  for each  $k$ . Lower semicontinuity of the functional  $\mathbf{f}$  implies

$$\hat{f}_\mathcal{B}^\mu(x_0) \leq \mathbf{f}(\mu_0^\varepsilon) + \varepsilon \leq \liminf_{k \rightarrow +\infty} \mathbf{f}(\mu_k) + \varepsilon \leq \lim_{k \rightarrow +\infty} \hat{f}_\mathcal{B}^\mu(x_{n_k}) + \varepsilon,$$

contradicting to (33) (since  $\varepsilon$  is arbitrary).

Lower semicontinuity of the concave lower bounded function  $\hat{f}_{\mathcal{B}}^{\mu}$  implies its  $\mu$ -concavity by Lemma 1.

B) The function  $\hat{f}_{\mathcal{B}}^{\sigma}$  is obviously well defined and  $\sigma$ -concave on the set  $\sigma\text{-co}(\mathcal{B})$ . Lower semicontinuity of this function is proved by a simple modification of the arguments of the proof of part A.

If  $\sigma\text{-co}(\mathcal{B}) = \overline{\text{co}}(\mathcal{B})$  then lower semicontinuity of the concave lower bounded function  $\hat{f}_{\mathcal{B}}^{\sigma}$  implies its  $\mu$ -concavity by Lemma 1. Since  $\hat{f}_{\mathcal{B}}^{\sigma}|_{\mathcal{B}} \geq f$  by the definition of  $\hat{f}_{\mathcal{B}}^{\sigma}$ , we have

$$\hat{f}_{\mathcal{B}}^{\sigma}(x) \geq \int_{\mathcal{B}} \hat{f}_{\mathcal{B}}^{\sigma}(y) \mu(dy) \geq \int_{\mathcal{B}} f(y) \mu(dy)$$

for any  $x \in \overline{\text{co}}(\mathcal{B})$  and any measure  $\mu$  in  $M_x(\mathcal{B})$ . This implies  $\hat{f}_{\mathcal{B}}^{\sigma} \geq \hat{f}_{\mathcal{B}}^{\mu}$  and hence  $\hat{f}_{\mathcal{B}}^{\sigma} = \hat{f}_{\mathcal{B}}^{\mu}$ .  $\square$

**The proof of Proposition 3.** We divide the proof into two steps.

1) Prove that for an arbitrary finitely supported measure  $\mu_0 = \sum_{i=1}^m \pi_i \delta(x_i)$ , where  $\{x_i\}_{i=1}^m \subset \mathcal{A}_k$ ,  $\{\pi_i\}_{i=1}^m \in \mathfrak{P}_m$ ,  $m \in \mathbb{N}$ , and an arbitrary sequence  $\{x^n\} \subset \mathcal{A}$  converging to  $x^0 = \sum_{i=1}^m \pi_i x_i$  there exist a subsequence  $\{x^{n_k}\}$  and a sequence  $\{\mu_k\} \subset M^a(\mathcal{A}_k)$  such that  $\lim_k \mu_k = \mu_0$  and  $\mathbf{b}(\mu_k) = x^{n_k}$  for all  $k$ .

For  $k = 1$  the above assertion follows from openness of the map  $M^a(\mathcal{A}_1) \ni \mu \mapsto \mathbf{b}(\mu)$ . Assume this assertion holds for some particular  $k$  and deduce its validity for  $k + 1$ .

Let  $\mu_0 = \sum_{i=1}^m \pi_i \delta(x_i) \in M^a(\mathcal{A}_{k+1})$ , where  $\pi_i > 0$  for all  $i$  and  $\{x_i\}_{i=1}^m \not\subset \mathcal{A}_k$ , let  $\{x^n\}$  be a sequence converging to  $x^0 = \sum_{i=1}^m \pi_i x_i$ . For each  $i = \overline{1, m}$  we have  $x_i = \alpha_i y_i + (1 - \alpha_i) z_i$ , where  $y_i \in \mathcal{A}_k$ ,  $z_i \in \mathcal{A}_1$  and  $\alpha_i \in [0, 1]$ . Hence  $x^0 = \eta y^0 + (1 - \eta) z^0$ , where

$$\eta = \sum_{i=1}^m \alpha_i \pi_i \in (0, 1), \quad y^0 = \eta^{-1} \sum_{i=1}^m \alpha_i \pi_i y_i \in \mathcal{A}, \quad z^0 = (1 - \eta)^{-1} \sum_{i=1}^m (1 - \alpha_i) \pi_i z_i \in \mathcal{A}.$$

By stability of the set  $\mathcal{A}$  we may assume (by replacing the sequence  $\{x^n\}$  by some its subsequence) existence of sequences  $\{y^n\} \subset \mathcal{A}$  and  $\{z^n\} \subset \mathcal{A}$  converging respectively to  $y^0$  and  $z^0$  such that  $x^n = \eta y^n + (1 - \eta) z^n$ . By induction we may consider (again by passing to a subsequence) that there exist sequences  $\{\nu_n\} \subset M^a(\mathcal{A}_k)$  and  $\{\zeta_n\} \subset M^a(\mathcal{A}_1)$  converging to the measures

$$\nu_0 \doteq \eta^{-1} \sum_{i=1}^m \alpha_i \pi_i \delta(y_i) \quad \text{and} \quad \zeta_0 \doteq (1 - \eta)^{-1} \sum_{i=1}^m (1 - \alpha_i) \pi_i \delta(z_i)$$

correspondingly, such that  $\mathbf{b}(\nu_n) = y^n$  and  $\mathbf{b}(\zeta_n) = z^n$  for all  $n$ .

By definition of the weak convergence for arbitrary  $N$  and for arbitrary sufficiently small<sup>11</sup>  $\varepsilon > 0$  and  $\delta > 0$  there exists such  $\bar{n} > N$  that

$$\nu_{\bar{n}} = \sum_{i=1}^m \nu_{\bar{n}}^i + \nu_{\bar{n}}^r \quad \text{and} \quad \zeta_{\bar{n}} = \sum_{i=1}^m \zeta_{\bar{n}}^i + \zeta_{\bar{n}}^r, \quad (34)$$

<sup>11</sup>In what follows it is assumed that  $\varepsilon < 1/4$  and  $\delta$  is so small that  $\delta$ -vicinities of different points of the sets  $\{y_i\}_{i=1}^m$  and  $\{z_i\}_{i=1}^m$  do not intersect each other.

where  $\nu_{\bar{n}}^i$  and  $\zeta_{\bar{n}}^i$  are measures with finite support contained respectively in  $U_\delta(y_i)$  and in  $U_\delta(z_i)$  such that

$$|\nu_{\bar{n}}^i(U_\delta(y_i)) - \eta^{-1}\alpha_i\pi_i| < \eta^{-1}\varepsilon\pi_i, \quad |\zeta_{\bar{n}}^i(U_\delta(z_i)) - (1-\eta)^{-1}(1-\alpha_i)\pi_i| < (1-\eta)^{-1}\varepsilon\pi_i, \quad (35)$$

all atoms of the measures  $\nu_{\bar{n}}^i$  and  $\zeta_{\bar{n}}^i$  have rational weights,  $i = \overline{1, m}$ , and

$$\nu_{\bar{n}}^r(\mathcal{A}) < \eta^{-1}\varepsilon, \quad \zeta_{\bar{n}}^r(\mathcal{A}) < (1-\eta)^{-1}\varepsilon. \quad (36)$$

Existence of representation (34) is obvious if the sets  $\{y_i\}_{i=1}^m$  and  $\{z_i\}_{i=1}^m$  consist of different elements. If these sets contain coinciding elements existence of this representation can be shown by "splitting" atoms of the measures  $\nu_{\bar{n}}$  and  $\zeta_{\bar{n}}$  as follows. Suppose, for example,  $y_1 = y_2 = \dots = y_p = y$ . Then the component  $\sum_t \lambda_t \delta(y_t)$  of the measure  $\nu_{\bar{n}}$  having atoms within  $U_\delta(y)$  can be "decomposed" as

$$\sum_t \lambda_t \delta(y_t) = \sum_t \gamma_1 \lambda_t \delta(y_t) + \dots + \sum_t \gamma_p \lambda_t \delta(y_t),$$

where  $\gamma_i = \alpha_i \pi_i / (\alpha_1 \pi_1 + \dots + \alpha_p \pi_p)$ , and the measure  $\nu_{\bar{n}}^i$  is constructed by using the measure  $\gamma_i \sum_t \lambda_t \delta(y_t)$ .

For given  $i$  let  $\nu_{\bar{n}}^i = \sum_{j=1}^{n_i^y} \frac{p_{ij}^y}{q_i} \delta(y_{ij})$  and  $\zeta_{\bar{n}}^i = \sum_{j=1}^{n_i^z} \frac{p_{ij}^z}{q_i} \delta(z_{ij})$ , where  $p_*$  and  $q_*$  are natural numbers. One can find such natural numbers  $P_i$ ,  $Q_i^y$  and  $Q_i^z$  that  $\sum_{j=1}^{n_i^y} \frac{p_{ij}^y}{q_i} = \frac{P_i}{Q_i^y}$  and  $\sum_{j=1}^{n_i^z} \frac{p_{ij}^z}{q_i} = \frac{P_i}{Q_i^z}$ . Let  $d_i^y = (q_i Q_i^y)^{-1}$  and  $d_i^z = (q_i Q_i^z)^{-1}$ . By using the "decomposition"

$$\frac{p_{ij}^y}{q_i} \delta(y_{ij}) = \underbrace{d_i^y \delta(y_{ij}) + \dots + d_i^y \delta(y_{ij})}_{p_{ij}^y Q_i^y \text{ summands}}$$

we obtain the representation  $\nu_{\bar{n}}^i = \sum_{l=1}^{P_i q_i} d_i^y \delta(\bar{y}_i^l)$ , where  $\{\bar{y}_i^l\}_l$  is a set of  $P_i q_i$  elements (which may be coinciding) contained in  $U_\delta(y_i)$ . In the similar way we obtain the representation  $\zeta_{\bar{n}}^i = \sum_{l=1}^{P_i q_i} d_i^z \delta(\bar{z}_i^l)$ , where  $\{\bar{z}_i^l\}_l$  is a set of  $P_i q_i$  elements contained in  $U_\delta(z_i)$ .

Let

$$\begin{aligned} \mu_{\bar{n}} &= \eta \nu_{\bar{n}} + (1-\eta) \zeta_{\bar{n}} = \sum_{i=1}^m (\eta \nu_{\bar{n}}^i + (1-\eta) \zeta_{\bar{n}}^i) + \eta \nu_{\bar{n}}^r + (1-\eta) \zeta_{\bar{n}}^r \\ &= \sum_{i=1}^m \sum_{l=1}^{P_i q_i} (\eta d_i^y \delta(\bar{y}_i^l) + (1-\eta) d_i^z \delta(\bar{z}_i^l)) + \eta \nu_{\bar{n}}^r + (1-\eta) \zeta_{\bar{n}}^r \end{aligned}$$

be a measure with the barycenter  $\eta y_{\bar{n}} + (1-\eta) z_{\bar{n}} = x_{\bar{n}}$ . The measure

$$\hat{\mu}_{\bar{n}} = \sum_{i=1}^m \sum_{l=1}^{P_i q_i} (\eta d_i^y + (1-\eta) d_i^z) \delta(\bar{x}_i^l) + \eta \nu_{\bar{n}}^r + (1-\eta) \zeta_{\bar{n}}^r, \quad \text{where } \bar{x}_i^l = \frac{\eta d_i^y \bar{y}_i^l + (1-\eta) d_i^z \bar{z}_i^l}{\eta d_i^y + (1-\eta) d_i^z}$$

has the same barycenter and lies in  $M^a(\mathcal{A}_{k+1})$ . Since

$$\bar{\alpha}_i = \frac{\eta d_i^y}{\eta d_i^y + (1-\eta) d_i^z} = \frac{\frac{\eta P_i}{Q_i^y \pi_i}}{\frac{\eta P_i}{Q_i^y \pi_i} + \frac{(1-\eta) P_i}{Q_i^z \pi_i}}$$



and (35) implies  $|\frac{\eta P_i}{Q_i^y \pi_i} - \alpha_i| < \varepsilon$ ,  $|\frac{(1-\eta)P_i}{Q_i^z \pi_i} - (1 - \alpha_i)| < \varepsilon$ , it is easy to show that  $|\bar{\alpha}_i - \alpha_i| < 6\varepsilon$ . Thus we conclude that  $\bar{x}_i^l = \bar{\alpha} \bar{y}_i^l + (1 - \bar{\alpha}) \bar{z}_i^l \in U_{\delta(i)}(x_i)$  for all  $i = \overline{1, m}$  and  $l = \overline{1, P_i q_i}$ , where  $\delta(i) = 2\delta + C_{y_i, z_i}(6\varepsilon)$  (see Remark 1). Since  $P_i q_i (\eta d_i^y + (1 - \eta) d_i^z) = \eta \frac{P_i}{Q_i^y} + (1 - \eta) \frac{P_i}{Q_i^z}$ , by using (35) and (36) it is easy to show that

$$|\hat{\mu}_{\bar{n}}(U_{\delta(i)}(x_i)) - \pi_i| \leq 4\varepsilon \quad (37)$$

provided  $U_{\delta(i)}(x_i) \cap U_{\delta(i')}(x_{i'}) = \emptyset$  for all  $i \neq i'$ .

For natural  $l$  let  $n_l = \bar{n}$  and  $\mu_l = \hat{\mu}_{\bar{n}}$ , where  $\bar{n}$  and  $\hat{\mu}_{\bar{n}}$  are produced by the above construction with  $N = l$  and  $\varepsilon = \delta = 1/l$ . Then  $\mathbf{b}(\mu_l) = x^{n_l}$  and (37) implies weak convergence of the sequence  $\{\mu_l\}$  to the measure  $\mu_0$ .

2) Let  $\mu_0 = \sum_{i=1}^{+\infty} \pi_i \delta(x_i)$  be an arbitrary measure in  $M^a(\mathcal{A}_k)$  and  $\{x^n\} \subset \mathcal{A}$  be a sequence converging to  $x^0 = \sum_{i=1}^{+\infty} \pi_i x_i$ . For natural  $m$  let  $\mu_0^m = (\lambda_m)^{-1} \sum_{i=1}^m \pi_i \delta(x_i)$ , where  $\lambda_m = \sum_{i=1}^m \pi_i$ , and let  $\nu_0^m$  be a measure in  $M^a(\mathcal{A}_1)$  such that  $\mathbf{b}(\nu_0^m) = (1 - \lambda_m)^{-1} \sum_{i>m} \pi_i x_i$ .

Since the sequence  $\{\mu_0^m\}_m$  converges to the measure  $\mu_0$ , for given natural  $l$  there exists  $m_l$  such that  $\mu_0^{m_l} \in U_{1/l}(\mu_0)$  and  $\lambda_{m_l} > 1 - 1/l$ .<sup>12</sup> We have  $x^0 = \lambda_{m_l} \mathbf{b}(\mu_0^{m_l}) + (1 - \lambda_{m_l}) \mathbf{b}(\nu_0^{m_l})$ . By stability of the set  $\mathcal{A}$  we may assume (by replacing the sequence  $\{x^n\}$  by some its subsequence) existence of sequences  $\{y^n\} \subset \mathcal{A}$  and  $\{z^n\} \subset \mathcal{A}$  converging respectively to  $\mathbf{b}(\mu_0^{m_l})$  and  $\mathbf{b}(\nu_0^{m_l})$  such that  $x^n = \lambda_{m_l} y^n + (1 - \lambda_{m_l}) z^n$ .

By the first part of the proof we may consider (again by passing to a subsequence) that there exists a sequence  $\{\mu_n\} \subset M^a(\mathcal{A}_k)$  converging to the measure  $\mu_0^{m_l}$  such that  $\mathbf{b}(\mu_n) = y_n$  for all  $n$ . Hence there exists  $n_l > l$  such that  $\mu_{n_l} \in U_{1/l}(\mu_0^{m_l}) \subset U_{2/l}(\mu_0)$ . Let

$$\bar{\mu}_l = \lambda_{m_l} \mu_{n_l} + (1 - \lambda_{m_l}) \nu_{n_l},$$

where  $\nu_{n_l}$  is an arbitrary measure in  $M^a(\mathcal{A}_1)$  such that  $\mathbf{b}(\nu_{n_l}) = z^{n_l}$ .

It is easy to see that the sequence  $\{\bar{\mu}_l\}$  is contained in  $M^a(\mathcal{A}_k)$  and converges to the measure  $\mu_0$  while by the construction  $\mathbf{b}(\bar{\mu}_l) = \lambda_{m_l} y^{n_l} + (1 - \lambda_{m_l}) z^{n_l} = x^{n_l}$  for each  $l$ .  $\square$

I am grateful to A.S.Holevo and the participants of his seminar for useful discussion. I am also grateful to the organizers of the workshop *Thematic Program on Mathematics in Quantum Information* at the Fields Institute, where some of this work was done.

## Bibliography

- [1] C.Adami, N.J.Cerf, "Capacity of noisy quantum channel", Phys.Rev. A. V.56 P.3470-3485. 1997; arXiv: quant-ph/9609024.
- [2] P.M.Alberti, A.Uhlmann, "Stochasticity and partial order. Doubly stochastic maps and unitary mixing", VEB Deutscher Verlag Wiss., Berlin, 1981.
- [3] C.D.Aliprantis, K.C.Border, "Infinite dimensional analysis", Springer Verlag, 2006.

<sup>12</sup>The set  $M(\mathcal{A})$  can be considered as a metric space [23].

- [4] E.Alfsen, "Compact convex sets and boundary integrals", Springer Verlag, 1971.
- [5] C.H.Bennett C.H., P.W.Shor, J.A.Smolin, A.V.Thapliyal, "Entanglement-assisted capacity and the reverse Shannon theorem", arXiv: quant-ph/0106052.
- [6] R.Bhatia, "Matrix Analysis", Springer Verlag, 1997.
- [7] P.Billingsley, "Convergence of probability measures", John Willey and Sons. Inc., New York-London-Sydney-Toronto, 1968.
- [8] O.Bratteli, D.W.Robinson, "Operators algebras and quantum statistical mechanics", Springer Verlag, New York-Heidelberg-Berlin, vol.I, 1979.
- [9] R.O'Brien, "On the openness of the barycentre map", Math. Ann., V.223, N.3, P.207-212, 1976.
- [10] R.Grzaslewicz, "Extreme continuous function property", Acta.Math.Hungar., V.74, 93-99, 1997.
- [11] G.A.Edgar, "On the Radon-Nikodim property and martingale convergence", Lecture Notes in Mathematics, 645, P.62-76, 1978.
- [12] L.Q.Eifler, "Open mapping theorems for probability measures on metric spaces", Pacific Journal of Mathematics, V.66, P.89-97, 1976.
- [13] J.Eisert, M.M.Wolf, "Gaussian quantum channels", arXiv:quant-ph/0505151.
- [14] A.S.Holevo, "Statistical structure of quantum theory", Springer Verlag, 2001.
- [15] A.S.Holevo, "On complementary channels and the additivity problem", Probability Theory and Applications, V.51, N.1, P.134-143, 2006; arXiv: quant-ph/0509101.
- [16] A.S.Holevo, M.E.Shirokov, "Mutual and coherent informations for infinite-dimensional quantum channels", arXiv:1004.2495 [math-ph].
- [17] A.D.Joffe, W.M.Tikhomirov, "Theory of extremum problems", AP, NY, 1979.
- [18] S.Kullback, "Information theory and statistics", John Wiley and Sons, NY, 1959.
- [19] D.Leung, G.Smith, "'Continuity of quantum channel capacities", Commun. Math. Phys., V.292, P.201-215, 2009; arXiv:0810.4931 [quant-ph].
- [20] G.Lindblad, "Expectation and Entropy Inequalities for Finite Quantum Systems", Commun. Math. Phys., 1974. V.39. N.2. P.111-119.
- [21] M.A.Nielsen, I.L.Chuang, "Quantum Computation and Quantum Information", Cambridge University Press, 2000.

- [22] S.Papadopoulou, "On the geometry of stable compact convex sets", *Math. Ann.*, V.229, P.193-200, 1977.
- [23] K.Parthasarathy, "Probability measures on metric spaces", Academic Press, New York and London, 1967.
- [24] V.Yu.Protasov, M.E.Shirokov "Generalized compactness in linear spaces and its applications", *Sbornik:Mathematics*, V.200, N.5, P.697-722, 2009; arXiv:1002.3610 [math-ph].
- [25] R.Rockafellar, "Convex analysis", Tyrrell, 1970.
- [26] M.E.Shirokov, "On properties of quantum channels related to the classical capacity", *Theory of Probability and its Applications*, V.52, N.2, P.250-276, 2008.
- [27] M.E.Shirokov, "Continuity of the von Neumann entropy", *Commun. Math. Phys.*, 2010, V.296, N.3, P.625-654; arXiv: math-ph/0904.1963.
- [28] M.E.Shirokov, "On properties of the space of quantum states and their application to construction of entanglement monotones", *Izvestiya: Mathematics*, V.74, N.4, 2010; arXiv:0804.1515 [math-ph].
- [29] N.N.Vahania, V.I.Tarieladze, "Covariant operators of probability measures in locally convex spaces", *Theory of Probability and its Applications*, V.23, N.1, P.1-23, 1978.
- [30] J.Vesterstrom, "On open maps, compact convex sets and operator algebras", *J. London Math. Soc.*, V.6, N.2, P.289-297, 1973.
- [31] A.Wehr, "General properties of entropy", *Rev. Mod. Phys.*, V.50, P.221-250, 1978.
- [32] A.Wehr, "How chaotic is a state of a quantum system", *Rep. Math. Phys.*, V.6, P.15-28.