Commun. Math. Phys. 296, 625–654 (2010)
Digital Object Identifier (DOI) 10.1007/s00220-010-1007-x

Communications in
**Mathematical
Physics**

Continuity of the von Neumann Entropy

M. E. Shirokov

Steklov Mathematical Institute, Gubkina Str. 8, 11991 Moscow,
Russia. E-mail: msh@mi.ras.ru

Received: 20 April 2009 / Accepted: 10 November 2009
Published online: 2 March 2010 –

Abstract: A general method for proving continuity of the von Neumann entropy on subsets of positive trace-class operators is considered. This makes it possible to re-derive the known conditions for continuity of the entropy in more general forms and to obtain several new conditions. The method is based on a particular approximation of the von Neumann entropy by an increasing sequence of concave continuous unitary invariant functions defined using decompositions into finite rank operators. The existence of this approximation is a corollary of a general property of the set of quantum states as a convex topological space called the strong stability property. This is considered in the first part of the paper.

Contents

1. Introduction

The set of quantum states – density operators in a separable Hilbert space – plays the central role in analysis of general infinite dimensional quantum systems. One of the technical problems in this analysis is related to noncompactness of the set of quantum states and nonexistence of inner points of this set considered as a closed convex subset of the separable Banach space of all trace-class operators. Another technical problem consists in discontinuity and unboundedness of basic characteristics of quantum states such as the von Neumann entropy, the relative entropy, etc. The above problems can be partially overcome by using two special properties of the set of quantum states considered in detail in the first part of [25]. The first of them can be considered as a kind of "weak compactness" since it provides generalization to the set of quantum states of several results well known for compact convex sets (see Sect. 2) while the second one called the stability property reveals the special relation between the topology and the convex structure of the set of quantum states (see Subsect. 3.1). These two properties provide the foundation of analysis of continuity of several important characteristics of quantum systems and quantum channels (see [25] and the references therein).

In this paper we prove a stronger version of the stability property of the set of quantum states naturally called *strong stability* and consider its applications concerning the problem of approximation of concave (convex) functions on the set of quantum states and providing a new approach to analysis of continuity of such functions.

The main application of the strong stability property considered in this paper is the development of a method of proving continuity of the von Neumann entropy. In infinite dimensions the von Neumann entropy is a nonnegative concave lower semicontinuous function on the set of quantum states taking the value $+\infty$ on a dense subset of this set.[1] Nevertheless the von Neumann entropy has continuous bounded restrictions to some important subsets of quantum states, for example, to the set of states of the system of quantum oscillators with bounded mean energy. Since continuity of the entropy is a very desirable property in analysis of quantum systems, various sufficient continuity conditions have been obtained up to now. The earliest among them seems to be Simon's dominated convergence theorems presented in [26] and widely used in applications (the generalized forms of these theorems are presented in Corollary 4). Another useful continuity condition originally appeared in [29] (as far as I know) and can be formulated as continuity of the entropy on each subset of states characterized by bounded mean value of a given positive unbounded operator with discrete spectrum provided that its sequence of eigenvalues has a sufficient rate of increase (see Example 1). Some special conditions of continuity of the von Neumann entropy are considered in [24]. It turns out that the strong stability property of the set of quantum states (more precisely, the approximation technique based on this property) provides a new method of proving continuity of the von Neumann entropy on a set of quantum states based on the established relation between this property and the special *uniform approximation property* of this set defined via the relative entropy. Well known results concerning the relative entropy make it possible to show conservation of the uniform approximation property under different set-operations, which implies roughly speaking "preservation of continuity" of the entropy under these set-operations.

The proposed method makes it possible to re-derive the known conditions of continuity of the von Neumann entropy mentioned above (in the more general forms) and to

[1] Moreover, the set of states with finite entropy is a first category subset of the set of all quantum states [29].

obtain the several new (as far as I know) conditions which seems to be useful in analysis of quantum systems.

2. Preliminaries

Let $\mathcal{H}$ be a separable Hilbert space, $\mathfrak{B}(\mathcal{H})$ – the Banach space of all bounded operators in $\mathcal{H}$ with the operator norm $\|\cdot\|$, $\mathfrak{T}(\mathcal{H})$ – the Banach space of all trace-class operators in $\mathcal{H}$ with the trace norm $\|\cdot\|_1$, containing the cone $\mathfrak{T}_+(\mathcal{H})$ of all positive trace-class operators. The closed convex subsets

$$\mathfrak{T}_1(\mathcal{H}) = \{A \in \mathfrak{T}_+(\mathcal{H}) \mid \mathrm{Tr}A \leq 1\} \text{ and } \mathfrak{S}(\mathcal{H}) = \{A \in \mathfrak{T}_+(\mathcal{H}) \mid \mathrm{Tr}A = 1\}$$

are complete separable metric spaces with the metric defined by the trace norm. Operators in $\mathfrak{S}(\mathcal{H})$ are denoted $\rho, \sigma, \omega, \ldots$ and called density operators or states since each density operator uniquely defines a normal state on $\mathfrak{B}(\mathcal{H})$ [3].

In what follows $\mathcal{A}$ is a subset of the cone of positive trace-class operators.

We denote by $\mathrm{cl}(\mathcal{A})$, $\mathrm{co}(\mathcal{A})$, σ-$\mathrm{co}(\mathcal{A})$, $\overline{\mathrm{co}}(\mathcal{A})$ and $\mathrm{extr}(\mathcal{A})$ the closure, the convex hull, the σ-convex hull, the convex closure and the set of all extreme points of a set $\mathcal{A}$ correspondingly [12,22].

In what follows we consider functions on subsets of $\mathfrak{T}_+(\mathcal{H})$ taking values in $[-\infty, +\infty]$, which are *semibounded* (either lower or upper bounded) on these subsets.

We denote by $\mathrm{co}f$ and $\overline{\mathrm{co}}f$ the convex hull and the convex closure of a function f on a convex set $\mathcal{A}$ [12,22].

The set of all bounded continuous functions on a set $\mathcal{A}$ is denoted $C(\mathcal{A})$.

The set of all Borel probability measures on a closed set $\mathcal{A}$ endowed with the topology of weak convergence is denoted $\mathcal{P}(\mathcal{A})$. This set can be considered as a complete separable metric space [19]. The *barycenter* $\mathbf{b}(\mu)$ of the measure μ in $\mathcal{P}(\mathcal{A})$ is the operator in $\overline{\mathrm{co}}(\mathcal{A})$ defined by the Bochner integral

$$\mathbf{b}(\mu) = \int_{\mathcal{A}} A\mu(dA),$$

which always exists if the set $\mathcal{A}$ is bounded.

For arbitrary subset $\mathcal{B} \subseteq \overline{\mathrm{co}}(\mathcal{A})$ let $\mathcal{P}_{\mathcal{B}}(\mathcal{A})$ be the subset of $\mathcal{P}(\mathcal{A})$ consisting of all measures with barycenter in $\mathcal{B}$.

Let $\mathcal{P}^{\mathrm{a}}(\mathcal{A})$ be the subset of $\mathcal{P}(\mathcal{A})$ consisting of atomic measures and let $\mathcal{P}^{\mathrm{f}}(\mathcal{A})$ be the subset of $\mathcal{P}^{\mathrm{a}}(\mathcal{A})$ consisting of measures with a finite number of atoms. Each measure in $\mathcal{P}^{\mathrm{a}}(\mathcal{A})$ corresponds to a collection of operators $\{A_i\} \subset \mathcal{A}$ with probability distribution $\{\pi_i\}$ conventionally called an *ensemble* and denoted $\{\pi_i, A_i\}$. The barycenter of this measure is the average $\sum_i \pi_i A_i$ of the corresponding ensemble.

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

A semibounded function f on the set $\mathfrak{S}(\mathcal{H})$ is called *σ-concave* at a state $\rho_0 \in \mathfrak{S}(\mathcal{H})$ if the discrete Jensen's inequality

$$f(\rho_0) \geq \sum_i \pi_i f(\rho_i)$$

holds for an arbitrary *countable* ensemble $\{\pi_i, \rho_i\}$ of states in $\mathfrak{S}(\mathcal{H})$ with the average state ρ_0.

A semibounded universally measurable[2] function f on the set $\mathfrak{S}(\mathcal{H})$ is called *μ-concave* at a state $\rho_0 \in \mathfrak{S}(\mathcal{H})$ if the integral Jensen's inequality

$$f(\rho_0) \geq \int_{\mathfrak{S}(\mathcal{H})} f(\rho)\mu(d\rho)$$

holds for an arbitrary measure μ in $\mathcal{P}(\mathfrak{S}(\mathcal{H}))$ with the barycenter ρ_0.

σ-convexity and *μ-convexity* of a function f are naturally defined via the above notions applied to the function $-f$.

The examples of semibounded functions (in particular, Borel functions) on the set $\mathfrak{S}(\mathcal{H})$, which are convex but not σ-convex or σ-convex but not μ-convex at particular states as well as sufficient conditions for σ-convexity and μ-convexity of a convex function at any state are considered in [25].

The identity operator in a Hilbert space $\mathcal{H}$ and the identity transformation of the space $\mathfrak{T}(\mathcal{H})$ are denoted $I_{\mathcal{H}}$ and $\mathrm{Id}_{\mathcal{H}}$ correspondingly.

Following [11] an arbitrary positive unbounded operator in a Hilbert space with discrete spectrum of finite multiplicity is called the *$\mathfrak{H}$-operator*.

The set $\mathfrak{S}(\mathcal{H})$ is not compact if $\dim \mathcal{H} = +\infty$, but it has the property consisting in compactness of the pre-image $\mathbf{b}^{-1}(\mathcal{A}) \subset \mathcal{P}(\mathfrak{S}(\mathcal{H}))$ of any compact subset $\mathcal{A}$ of $\mathfrak{S}(\mathcal{H})$ under the map $\mu \mapsto \mathbf{b}(\mu)$ [11, Prop. 2], which can be used for proving for the set $\mathfrak{S}(\mathcal{H})$ and for its subsets several results well known for compact sets. This property (in the general context of a metrizable convex subset of a locally convex space) is studied in detail in [20], where it is called *μ-compactness*. It implies in particular the following Choquet-type assertion and the lemma below.

Lemma 1. *Let $\mathcal{A}$ be a closed subset of $\mathfrak{S}(\mathcal{H})$. For an arbitrary state ρ in $\overline{\mathrm{co}}(\mathcal{A})$ there exists a measure μ in $\mathcal{P}(\mathcal{A})$ such that $\boldsymbol{b}(\mu) = \rho$.*

Proof. Let $\rho_0 \in \overline{\mathrm{co}}(\mathcal{A})$ and $\{\rho_n\} \subset \mathrm{co}(\mathcal{A})$ be a sequence converging to the state ρ_0. For each $n \in \mathbb{N}$ there exists a measure $\mu_n \in \mathcal{P}(\mathcal{A})$ with finite support such that $\rho_n = \mathbf{b}(\mu_n)$. By μ-compactness of the set $\mathfrak{S}(\mathcal{H})$ the sequence $\{\mu_n\}$ has a partial limit $\mu_0 \in \mathcal{P}(\mathcal{A})$. Continuity of the map $\mu \mapsto \mathbf{b}(\mu)$ implies $\mathbf{b}(\mu_0) = \rho_0$. $\square$

Lemma 1 provides correctness of the definition of the functions in the following lemma, proved in the Appendix.

Lemma 2. *Let f be a lower semicontinuous lower bounded function on a closed subset $\mathcal{A}$ of $\mathfrak{S}(\mathcal{H})$.*

A) *The function*

$$\check{f}_{\mathcal{A}}(\rho) = \inf_{\mu \in \mathcal{P}_{\{\rho\}}(\mathcal{A})} \int_{\mathcal{A}} f(\sigma)\mu(d\sigma)$$

is convex and lower semicontinuous on the set $\overline{\mathrm{co}}(\mathcal{A})$. For arbitrary ρ in $\overline{\mathrm{co}}(\mathcal{A})$ the infimum in the definition of the value $\check{f}_{\mathcal{A}}(\rho)$ is achieved at a particular measure in $\mathcal{P}_{\{\rho\}}(\mathcal{A})$.

B) *If the map $\mathcal{P}(\mathcal{A}) \ni \mu \mapsto \boldsymbol{b}(\mu) \in \overline{\mathrm{co}}(\mathcal{A})$ is open then the function*

$$\hat{f}_{\mathcal{A}}(\rho) = \sup_{\mu \in \mathcal{P}_{\{\rho\}}(\mathcal{A})} \int_{\mathcal{A}} f(\sigma)\mu(d\sigma)$$

is concave and lower semicontinuous on the set $\overline{\mathrm{co}}(\mathcal{A})$.

[2] This means that the function f is measurable with respect to any measure in $\mathcal{P}(\mathfrak{S}(\mathcal{H}))$.

For given natural k we denote by $\mathfrak{T}_+^k(\mathcal{H})$ (correspondingly by $\mathfrak{S}_k(\mathcal{H})$) the set of positive trace-class operators (correspondingly states) having rank $\leq k$. The convex set $\bigcup_{k=1}^{+\infty}\mathfrak{S}_k(\mathcal{H})$ of all finite rank states is denoted $\mathfrak{S}_{\mathrm{f}}(\mathcal{H})$.

A linear positive trace-nonincreasing map $\Phi : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H})$ such that the dual map $\Phi^* : \mathfrak{B}(\mathcal{H}) \to \mathfrak{B}(\mathcal{H})$ is completely positive is called a *quantum operation* [10]. The convex set of all quantum operations from $\mathfrak{T}(\mathcal{H})$ to itself is denoted $\mathfrak{F}_{\leq 1}(\mathcal{H})$. If a quantum operation Φ is trace-preserving then it is called a *quantum channel*.

An arbitrary quantum operation (correspondingly channel) $\Phi \in \mathfrak{F}_{\leq 1}(\mathcal{H})$ has the following Kraus representation:

$$\Phi(\cdot) = \sum_{j=1}^{+\infty} V_j(\cdot)V_j^*,$$

where $\{V_j\}_{j=1}^{+\infty}$ is a set of operators in $\mathfrak{B}(\mathcal{H})$ such that $\sum_{j=1}^{+\infty} V_j^* V_j \leq I_{\mathcal{H}}$ (correspondingly $\sum_{j=1}^{+\infty} V_j^* V_j = I_{\mathcal{H}}$).

For given natural n we denote by $\mathfrak{F}_{\leq 1}^n(\mathcal{H})$ the subset of $\mathfrak{F}_{\leq 1}(\mathcal{H})$ consisting of quantum operations having the Kraus representation with $\leq n$ nonzero summands.

We will use the following result of the purification theory.

Lemma 3. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces such that* $\dim\mathcal{H} = \dim\mathcal{K}$. *For an arbitrary pure state ω_0 in $\mathfrak{S}(\mathcal{H}\otimes\mathcal{K})$ and an arbitrary sequence $\{\rho_n\}$ of states in $\mathfrak{S}(\mathcal{H})$ converging to the state $\rho_0 = \mathrm{Tr}_{\mathcal{K}}\omega_0$ there exists a sequence $\{\omega_n\}$ of pure states in $\mathfrak{S}(\mathcal{H}\otimes\mathcal{K})$ converging to the state ω_0 such that $\rho_n = \mathrm{Tr}_{\mathcal{K}}\omega_n$ for all n.*

The assertion of this lemma can be proved by noting that the infimum in the definition of the Bures distance (or the supremum in the definition of the Uhlmann fidelity) between two quantum states can be taken only over all purifications of one state with fixed purification of another state and that convergence of a sequence of states in the trace norm distance implies its convergence in the Bures distance [9,14].

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

Note. In what follows continuity of a function f on a set $\mathcal{A} \subset \mathfrak{T}_+(\mathcal{H})$ implies its finiteness on this set (in contrast to lower (upper) semicontinuity).

3. The Strong Stability Property of $\mathfrak{S}(\mathcal{H})$

3.1. The definition. The notion of stability of a convex subset of a linear topological space appeared at the end of the 1970's as a result of study of the properties of compact convex sets, which led in particular to proving equivalence of continuity of the convex hull (envelope) of an arbitrary continuous function (the CE-property), openness of the mixture map and openness of the barycenter map for given compact convex set (the Vesterstrom-O'Brien theorem [4]). In the subsequent papers (see [5,8,18] and the references therein) the term *stability* was used to denote openness of the mixture map for an arbitrary convex subset of a linear topological space (which is not equivalent in general to the CE-property).

The stability property of the set $\mathfrak{S}(\mathcal{H})$ of quantum states and its corollaries are considered in detail in [25]. It consists in the validity of the following equivalent[3] statements:

[3] Equivalence of these statements follows from the μ-compact generalization of the Vesterstrom-O'Brien theorem [20, Theorem 1].

- the map $\mathfrak{S}(\mathcal{H})^{\times 2} \times [0,1] \ni (\rho, \sigma, \lambda) \mapsto \lambda\rho + (1-\lambda)\sigma \in \mathfrak{S}(\mathcal{H})$ is open;
- the map $\mathcal{P}(\mathfrak{S}(\mathcal{H})) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathfrak{S}(\mathcal{H})$ is open;
- the map $\mathcal{P}(\mathrm{extr}\mathfrak{S}(\mathcal{H})) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathfrak{S}(\mathcal{H})$ is open;
- $\mathrm{co} f = \overline{\mathrm{co}} f \in C(\mathfrak{S}(\mathcal{H}))$ for arbitrary $f \in C(\mathfrak{S}(\mathcal{H}))$;
- $f_*^{\sigma} = f_*^{\mu} \in C(\mathfrak{S}(\mathcal{H}))$ for arbitrary $f \in C(\mathrm{extr}\mathfrak{S}(\mathcal{H}))$, where f_*^{σ} and f_*^{μ} are the σ-convex roof and the μ-convex roof of the function f [25].

Physically openness of the map $\mathcal{P}(\mathfrak{S}(\mathcal{H})) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathfrak{S}(\mathcal{H})$ (correspondingly of the map $\mathcal{P}(\mathrm{extr}\mathfrak{S}(\mathcal{H})) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathfrak{S}(\mathcal{H})$) means roughly speaking that any small perturbation of the average state of a given continuous ensemble of states (correspondingly of pure states) can be realized by appropriate small perturbations of the states of this ensemble.

It turns out that the stability property of the set $\mathfrak{S}(\mathcal{H})$ can be enforced by showing that any small perturbation of the average state of a given (countable or continuous) ensemble of finite rank states can be realized by appropriate small perturbations of the states of this ensemble *without increasing the maximal rank of these states*. Mathematically this *strong stability property* of the set $\mathfrak{S}(\mathcal{H})$ is formulated in the following theorem.

Theorem 1. *The surjective continuous maps $\mathcal{P}(\mathfrak{S}_k(\mathcal{H})) \ni \mu \mapsto \boldsymbol{b}(\mu) \in \mathfrak{S}(\mathcal{H})$ and $\mathcal{P}^{\mathrm{a}}(\mathfrak{S}_k(\mathcal{H})) \ni \mu \mapsto \boldsymbol{b}(\mu) \in \mathfrak{S}(\mathcal{H})$ are open for each natural k.*

As mentioned before, the assertion of Theorem 1 for $k = 1$ is equivalent to openness of the map $\mathcal{P}(\mathfrak{S}(\mathcal{H})) \ni \mu \mapsto \mathbf{b}(\mu) \in \mathfrak{S}(\mathcal{H})$. The proof of this equivalence is based on coincidence of the set $\mathfrak{S}_1(\mathcal{H})$ with the set $\mathrm{extr}\mathfrak{S}(\mathcal{H})$ and is universal in the sense that it is valid for an arbitrary compact or μ-compact convex set in the role of $\mathfrak{S}(\mathcal{H})$ [4,20]. In contrast to this in the proof of the assertion of Theorem 1 for $k > 1$ the specific structure of the set $\mathfrak{S}(\mathcal{H})$ is essentially used.

The basic ingredients of the proof of the above theorem are the following lemma and Lemma 5 below.

Lemma 4. *Let $\{\pi_i^0, \rho_i^0\}$ be a countable ensemble of states in $\mathfrak{S}_k(\mathcal{H})$ with the average $\rho_0 = \sum_{i=1}^{+\infty} \pi_i^0 \rho_i^0$. For an arbitrary sequence $\{\rho_n\} \subset \mathfrak{S}(\mathcal{H})$ converging to the state ρ_0 there exists a sequence $\{\{\pi_i^n, \rho_i^n\}\}_n$ of countable ensembles of states in $\mathfrak{S}_k(\mathcal{H})$ such that*

$$\lim_{n\to+\infty} \pi_i^n = \pi_i^0, \quad \pi_i^0 > 0 \Rightarrow \lim_{n\to+\infty} \rho_i^n = \rho_i^0, \quad \forall i, \quad \textit{and} \quad \rho_n = \sum_{i=1}^{+\infty} \pi_i^n \rho_i^n, \quad \forall n.$$

The assertion of this lemma implies weak convergence of the sequence $\{\{\pi_i^n, \rho_i^n\}\}_n$ of atomic measures to the atomic measure $\{\pi_i^0, \rho_i^0\}$, t.i. convergence in $\mathcal{P}(\mathfrak{S}_k(\mathcal{H}))$, which means that $\lim_{n\to+\infty} \sum_i \pi_i^n f(\rho_i^n) = \sum_i \pi_i^0 f(\rho_i^0)$ for any function f in $C(\mathfrak{S}_k(\mathcal{H}))$. This relation can be easily proved by noting that pointwise convergence of the sequence $\{\{\pi_i^n\}\}_n$ to the probability distribution $\{\pi_i^0\}$ implies its convergence in the norm of total variation.

Proof of Lemma 4. For each i let $|\varphi_i\rangle$ be a unit vector in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{H}_k)$ such that $\mathrm{Tr}_{\mathcal{H}_k} |\varphi_i\rangle\langle\varphi_i| = \rho_i^0$, where $\mathcal{H}_k$ is an auxiliary k-dimensional Hilbert space. Let $\{|\epsilon_i\rangle\}_{i=1}^{+\infty}$ be an orthonormal basis in a separable Hilbert space $\mathcal{H}'$. Consider the unit vector $|\psi_0\rangle = \sum_{i=1}^{+\infty} \sqrt{\pi_i^0} |\varphi_i\rangle \otimes |\epsilon_i\rangle$ in the space $\mathcal{H} \otimes \mathcal{H}_k \otimes \mathcal{H}'$. It is easy to see that $\mathrm{Tr}_{\mathcal{H}_k \otimes \mathcal{H}'} |\psi_0\rangle\langle\psi_0| = \rho_0$. By Lemma 3 there exists sequence $\{|\psi_n\rangle\}$ of unit vectors in $\mathcal{H} \otimes \mathcal{H}_k \otimes \mathcal{H}'$ converging to the vector $|\psi_0\rangle$ such that $\mathrm{Tr}_{\mathcal{H}_k \otimes \mathcal{H}'} |\psi_n\rangle\langle\psi_n| = \rho_n$ for each n.

Let $\{E_i = I_{\mathcal{H}} \otimes I_{\mathcal{H}_k} \otimes |\epsilon_i\rangle\langle\epsilon_i|\}_{i=1}^{+\infty}$ be the local measurement in the space $\mathcal{H}\otimes\mathcal{H}_k\otimes\mathcal{H}'$ [10]. Since $E_i|\psi_0\rangle = \sqrt{\pi_i^0}|\varphi_i\rangle \otimes |\epsilon_i\rangle$ for each i we have $\pi_i^0 = \mathrm{Tr}E_i|\psi_0\rangle\langle\psi_0|$ and $\pi_i^0\rho_i^0 = \mathrm{Tr}_{\mathcal{H}_k\otimes\mathcal{H}'}E_i|\psi_0\rangle\langle\psi_0|E_i$. Let $\pi_i^n = \mathrm{Tr}E_i|\psi_n\rangle\langle\psi_n|$ and

$$\rho_i^n = \begin{cases} (\pi_i^n)^{-1}\mathrm{Tr}_{\mathcal{H}_k\otimes\mathcal{H}'}E_i|\psi_n\rangle\langle\psi_n|E_i, & \pi_i^n > 0 \\ \rho_i^0, & \pi_i^n = 0. \end{cases}$$

Then $\mathrm{rank}\rho_i^n \leq k$ for all n and i. The sequence of ensembles $\{\pi_i^n, \rho_i^n\}$ has the required properties. $\square$

Remark 1. It is interesting to compare the above lemma with Lemma 3 in [23] containing the analogous assertion concerning finite ensembles with no rank restriction on states of ensembles. The case of the finite ensemble $\{\pi_i^0, \rho_i^0\}_{i=1}^m$ is naturally embedded in the condition of Lemma 4 by setting $\pi_i^0 = 0$ for all $i > m$, but this lemma does not guarantee that the sequence $\{\{\pi_i^n, \rho_i^n\}\}_n$ consists of ensembles of m states in contrast to the assertion of Lemma 3 in [23]. Increasing dimensionality of ensembles of the sequence $\{\{\pi_i^n, \rho_i^n\}\}_n$ is the cost of the rank restriction on the states of these ensembles.

For arbitrary state ρ in $\mathfrak{S}(\mathcal{H})$ the set $\mathcal{P}^{\mathrm{a}}_{\{\rho\}}(\mathfrak{S}(\mathcal{H}))$ is a dense subset of $\mathcal{P}_{\{\rho\}}(\mathfrak{S}(\mathcal{H}))$ [11, Lemma 1]. This simple result can be enforced as follows.

Lemma 5. *For arbitrary state ρ in $\mathfrak{S}(\mathcal{H})$ and $k \in \mathbb{N}$ the set $\mathcal{P}^{\mathrm{a}}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))$ is a dense subset of $\mathcal{P}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))$.*

This means that any probability measure supported by the set of states of rank $\leq k$ can be weakly approximated by some sequence of atomic measures – countable ensembles of states of rank $\leq k$ with the same barycenter.

Proof. To prove the assertion of the lemma for $k = 1$ consider the Choquet ordering on the set $\mathcal{P}(\mathfrak{S}(\mathcal{H}))$. We say that $\mu \succ \nu$ if and only if

$$\int_{\mathfrak{S}(\mathcal{H})} f(\sigma)\mu(d\sigma) \geq \int_{\mathfrak{S}(\mathcal{H})} f(\sigma)\nu(d\sigma)$$

for an arbitrary convex continuous bounded function f on the set $\mathfrak{S}(\mathcal{H})$ [7].

By Lemma 1 in [11] for given measure μ_0 in $\mathcal{P}(\mathfrak{S}_1(\mathcal{H}))$ there exists a sequence $\{\mu_n\}$ of measures in $\mathcal{P}(\mathfrak{S}(\mathcal{H}))$ with finite support converging to the measure μ_0 such that $\mathbf{b}(\mu_n) = \mathbf{b}(\mu_0)$ for all n. Decomposing each atom of the measure μ_n into a convex combination of pure states we obtain the measure $\hat{\mu}_n$ in $\mathcal{P}^{\mathrm{a}}(\mathfrak{S}_1(\mathcal{H}))$ with the same barycenter. It is easy to see that $\hat{\mu}_n \succ \mu_n$. By μ-compactness of the set $\mathfrak{S}(\mathcal{H})$ the set $\{\hat{\mu}_n\}_{n>0}$ is relatively compact. This implies existence of the subsequence $\{\hat{\mu}_{n_k}\}$ converging to a measure $\{\hat{\mu}_0\}$ in $\mathcal{P}(\mathfrak{S}_1(\mathcal{H}))$ [19, Theorem 6.1]. 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 μ_0 with respect to the Choquet ordering (see the Appendix). Density of the set $\mathcal{P}^{\mathrm{a}}_{\{\rho\}}(\mathfrak{S}_1(\mathcal{H}))$ in $\mathcal{P}_{\{\rho\}}(\mathfrak{S}_1(\mathcal{H}))$ is proved.

Let $k > 1$ and $\mathcal{H}_k$ be the k-dimensional Hilbert space. Let Π be the multi-valued map from $\mathfrak{S}_k(\mathcal{H})$ into the set $2^{\mathfrak{S}_1(\mathcal{H}\otimes\mathcal{H}_k)}$ such that $\Pi(\rho)$ is the set of all purifications in $\mathfrak{S}_1(\mathcal{H}\otimes\mathcal{H}_k)$ of the state $\rho \in \mathfrak{S}_k(\mathcal{H})$. It is clear that the map Π is closed-valued. Thus by Theorem 3.1 in [28] to prove existence of a measurable selection of the map Π it is sufficient to show weak measurability of this map in terms of [28]. Let U be an open subset of $\mathfrak{S}_1(\mathcal{H}\otimes\mathcal{H}_k)$. Then $\Pi^-(U)=\{\rho\in\mathfrak{S}_k(\mathcal{H})\,|\,\Pi(\rho)\cap U\neq\emptyset\}=\Theta(U)$, where $\Theta(\cdot)=\mathrm{Tr}_{\mathcal{H}_k}(\cdot)$ is the affine (single valued) map from $\mathfrak{S}(\mathcal{H}\otimes\mathcal{H}_k)$ onto $\mathfrak{S}_k(\mathcal{H})$. Since the restriction of the map Θ to the set $\mathfrak{S}_1(\mathcal{H}\otimes\mathcal{H}_k)$ is open,[4] the set $\Pi^-(U)=\Theta(U)$ is open and hence Borel. As mentioned before, this implies existence of a measurable selection Π_* of the map Π.

Let $\nu_0=\mu_0\circ\Pi_*^{-1}$ be the image of the measure μ_0 under the map Π_*. It is clear that $\nu_0\in\mathcal{P}(\mathfrak{S}_1(\mathcal{H}\otimes\mathcal{H}_k))$. By the assertion of the lemma for $k=1$ there exists a sequence $\{\nu_n\}$ of measures in $\mathcal{P}^{\mathrm{a}}(\mathfrak{S}_1(\mathcal{H}\otimes\mathcal{H}_k))$ converging to the measure ν_0 such that $\mathbf{b}(\nu_n)=\mathbf{b}(\nu_0)$ for all n. Since $\Theta\circ\Pi_*=\mathrm{Id}_{\mathcal{H}}$ the image $\nu_0\circ\Theta^{-1}$ of the measure ν_0 under the map Θ coincides with μ_0. This and continuity of the map Θ imply convergence of the sequence $\{\mu_n=\nu_n\circ\Theta^{-1}\}$ of measures in $\mathcal{P}^{\mathrm{a}}(\mathfrak{S}_k(\mathcal{H}))$ to the measure μ_0. Since the map Θ is affine we have

$$\mathbf{b}(\mu_n)=\Theta(\mathbf{b}(\nu_n))=\Theta(\mathbf{b}(\nu_0))=\mathbf{b}(\mu_0)$$

for all n. Thus the sequence $\{\mu_n\}$ has the required properties. □

Proof of Theorem 1. By Lemma 5 it is sufficient to prove openness of the surjective map $\mathcal{P}^{\mathrm{a}}(\mathfrak{S}_k(\mathcal{H}))\ni\mu\mapsto\mathbf{b}(\mu)\in\mathfrak{S}(\mathcal{H})$ for each natural k.

Let U be an arbitrary open subset of $\mathcal{P}^{\mathrm{a}}(\mathfrak{S}_k(\mathcal{H}))$. Suppose $\mathbf{b}(U)$ is not open in $\mathfrak{S}(\mathcal{H})$. Then there exist a state $\rho_0\in\mathbf{b}(U)$ and a sequence $\{\rho_n\}$ of states in $\mathfrak{S}(\mathcal{H})\setminus\mathbf{b}(U)$ converging to the state ρ_0.

Let $\mu_0=\{\pi_i^0,\rho_i^0\}$ be a measure in U such that $\mathbf{b}(\mu_0)=\rho_0$. By Lemma 4 (and the remark after it) there exists a sequence of measures $\mu_n=\{\pi_i^n,\rho_i^n\}$ in $\mathcal{P}^{\mathrm{a}}(\mathfrak{S}_k(\mathcal{H}))$ converging to the measure $\mu_0=\{\pi_i^0,\rho_i^0\}$ such that $\mathbf{b}(\mu_n)=\rho_n$ for all n. Openness of the set U implies $\mu_n\in U$ for all sufficiently large n contradicting the choice of the sequence $\{\rho_n\}$. □

3.2. Some implications. In the case $\dim\mathcal{H}<+\infty$ the convex (concave) roof extension to the set $\mathfrak{S}(\mathcal{H})$ of a function f on the set of pure states $\mathfrak{S}_1(\mathcal{H})=\mathrm{extr}\mathfrak{S}(\mathcal{H})$ is defined at a mixed state ρ as the minimal (maximal) value of $\sum_i\pi_i f(\rho_i)$ over all decompositions $\rho=\sum_i\pi_i\rho_i$ of this state into a finite convex combination of pure states [27]. This extension is widely used in quantum information theory, in particular, in construction of entanglement monotones [21]. The convex (concave) roof extension has the two natural generalizations to the case $\dim\mathcal{H}=+\infty$, called in [25] the σ-convex (concave) roof and the μ-convex (concave) roof correspondingly (the first extension is defined via all decompositions of a state into a countable convex combination of pure states while the second one – via all "continuous" decompositions corresponding to Borel probability measures on the set of pure states with given barycenter).

Generalizing the σ-concave roof construction, for given natural k and semibounded function f on the set $\mathfrak{S}_k(\mathcal{H})$ consider the function

$$\mathfrak{S}(\mathcal{H})\ni\rho\mapsto\hat{f}_k^{\sigma}(\rho)=\sup_{\{\pi_i,\rho_i\}\in\mathcal{P}^{\mathrm{a}}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))}\sum_i\pi_i f(\rho_i)$$

[4] This means that for an arbitrary state $\omega_0\in\mathfrak{S}_1(\mathcal{H}\otimes\mathcal{H}_k)$ and sequence $\{\rho_n\}\subset\mathfrak{S}_k(\mathcal{H})$ converging to the state $\rho_0=\Theta(\omega_0)$ there exist a subsequence $\{\rho_{n_k}\}$ and a sequence $\{\omega_k\}\subset\mathfrak{S}_1(\mathcal{H}\otimes\mathcal{H}_k)$ converging to the state ω_0 such that $\Theta(\omega_k)=\rho_{n_k}$ for all k. The last property can be verified by using the standard arguments of the purification theory.

(the supremum is over all decompositions of the state ρ into a countable convex combination of states of rank $\leq k$). This function is obviously σ-concave on the set $\mathfrak{S}(\mathcal{H})$ (see Sect. 2). If the function f is σ-concave at any state in $\mathfrak{S}_k(\mathcal{H})$ then the functions $\hat{f}_k^\sigma$ and f coincide on the set $\mathfrak{S}_k(\mathcal{H})$, so in this case the function $\hat{f}_k^\sigma$ can be considered as an extension of the function f to the set $\mathfrak{S}(\mathcal{H})$.

Generalizing the μ-concave roof construction, for given natural k and semibounded Borel function f on the set $\mathfrak{S}_k(\mathcal{H})$ consider the function

$$\mathfrak{S}(\mathcal{H}) \ni \rho \mapsto \hat{f}_k^\mu(\rho) = \sup_{\mu \in \mathcal{P}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))} \int_{\mathfrak{S}_k(\mathcal{H})} f(\sigma)\mu(d\sigma)$$

(the supremum is over all probability measures with the barycenter ρ supported by states of rank $\leq k$). This function is also obviously σ-concave on the set $\mathfrak{S}(\mathcal{H})$ but its μ-concavity depends on the question of its universal measurability. By Propositions 1 and 2 below, the function $\hat{f}_k^\mu$ is μ-concave on the set $\mathfrak{S}(\mathcal{H})$ if the function f is either lower bounded lower semicontinuous or upper bounded upper semicontinuous on the set $\mathfrak{S}_k(\mathcal{H})$. If the function f is μ-concave at any state in $\mathfrak{S}_k(\mathcal{H})$ then the functions $\hat{f}_k^\mu$ and f coincide on the set $\mathfrak{S}_k(\mathcal{H})$, so in this case the function $\hat{f}_k^\mu$ can be considered as an extension of the function f to the set $\mathfrak{S}(\mathcal{H})$.

The strong stability property of the set $\mathfrak{S}(\mathcal{H})$ stated in Theorem 1 and Lemma 5 imply the following result.

Proposition 1. *Let f be a lower semicontinuous lower bounded function on the set $\mathfrak{S}_k(\mathcal{H})$. Then $\hat{f}_k^\sigma = \hat{f}_k^\mu$ and this function is lower semicontinuous and μ-concave on the set $\mathfrak{S}(\mathcal{H})$.*

Proof. Coincidence of the functions $\hat{f}_k^\sigma$ and $\hat{f}_k^\mu$ follows from lower semicontinuity of the functional $\mathcal{P}(\mathfrak{S}_k(\mathcal{H})) \ni \mu \mapsto \int_{\mathfrak{S}_k(\mathcal{H})} f(\sigma)\mu(d\sigma)$ (proved by the standard argumentation) and Lemma 5. Theorem 1 and Lemma 2 imply lower semicontinuity of the lower bounded function $\hat{f}_k^\sigma = \hat{f}_k^\mu$, which guarantees its μ-concavity (by Proposition A-2 in the Appendix in [25]). $\square$

The μ-compactness property of the set $\mathfrak{S}(\mathcal{H})$ (described before Lemma 1) implies the following result.

Proposition 2. *Let f be an upper semicontinuous upper bounded function on the set $\mathfrak{S}_k(\mathcal{H})$. Then the function $\hat{f}_k^\mu$ is upper semicontinuous and μ-concave on the set $\mathfrak{S}(\mathcal{H})$.*

For an arbitrary state ρ in $\mathfrak{S}(\mathcal{H})$ the supremum in the definition of the value $\hat{f}_k^\mu(\rho)$ is achieved at some measure in $\mathcal{P}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))$.

Proof. Lemma 2 implies attainability of the supremum in the definition of the value $\hat{f}_k^\mu(\rho)$ and upper semicontinuity of the function $\hat{f}_k^\mu$, which guarantees its μ-concavity (by Proposition A-2 in the Appendix in [25]). $\square$

Under the condition of Proposition 2 we can say nothing about upper semicontinuity and μ-concavity of the function $\hat{f}_k^\sigma$ (see example 2 in [25]).

The above two propositions have the obvious corollary.

Corollary 1. *Let f be a continuous bounded function on the set $\mathfrak{S}_k(\mathcal{H})$. Then $\hat{f}_k^\sigma = \hat{f}_k^\mu$ and this function is continuous on the set $\mathfrak{S}(\mathcal{H})$.*

4. On Approximation of Concave (Convex) Functions on $\mathfrak{S}(\mathcal{H})$

The functional constructions considered in Subsect. 3.2 can be used in the study of the following *approximation problem*: for a given concave (convex) function f on the set $\mathfrak{S}(\mathcal{H})$ having some particular symmetry[5] to find a monotonic sequence $\{f_k\}$ of concave (convex) functions on the set $\mathfrak{S}(\mathcal{H})$ having the same symmetry, satisfying additional analytical requirements and such that

$$f_k|_{\mathfrak{S}_k(\mathcal{H})} = f|_{\mathfrak{S}_k(\mathcal{H})}, \ \forall k, \quad \text{and} \quad \lim_{k\to+\infty} f_k(\rho) = f(\rho), \ \forall \rho \in \mathfrak{S}(\mathcal{H}).$$

Let f be a function on the set $\mathfrak{S}(\mathcal{H})$ having semibounded restriction to the set $\mathfrak{S}_k(\mathcal{H})$ for each k. We can consider the nondecreasing sequence $\{\hat{f}_k^\sigma\}$ of concave functions on the set $\mathfrak{S}(\mathcal{H})$ and its pointwise limit $\hat{f}_{+\infty}^\sigma = \sup_k \hat{f}_k^\sigma$. If the restriction of the function f to the set $\mathfrak{S}_k(\mathcal{H})$ is universally measurable for each k then we can also consider the nondecreasing sequence $\{\hat{f}_k^\mu\}$ of concave functions on the set $\mathfrak{S}(\mathcal{H})$ and its pointwise limit $\hat{f}_{+\infty}^\mu = \sup_k \hat{f}_k^\mu$.

By construction all the functions in the sequences $\{\hat{f}_k^\sigma\}$ and $\{\hat{f}_k^\mu\}$ inherit the arbitrary symmetry of the function f. Hence the same assertion holds for the functions $\hat{f}_{+\infty}^\sigma$ and $\hat{f}_{+\infty}^\mu$.

The functions $\hat{f}_{+\infty}^\sigma$ and $\hat{f}_{+\infty}^\mu$ are concave on the set $\mathfrak{S}(\mathcal{H})$. By construction $\hat{f}_{+\infty}^\sigma \leq \hat{f}_{+\infty}^\mu$ and $f|_{\mathfrak{S}_\mathrm{f}(\mathcal{H})} \leq \hat{f}_{+\infty}^\sigma|_{\mathfrak{S}_\mathrm{f}(\mathcal{H})}$ ($\mathfrak{S}_\mathrm{f}(\mathcal{H})$ is the convex subset of $\mathfrak{S}(\mathcal{H})$ consisting of finite rank states). If the function f is σ-concave on the set $\mathfrak{S}(\mathcal{H})$ then $\hat{f}_{+\infty}^\sigma \leq f$, if the function f is μ-concave on the set $\mathfrak{S}(\mathcal{H})$ then $\hat{f}_{+\infty}^\mu \leq f$. To show coincidence of the functions $\hat{f}_{+\infty}^\sigma$ and $\hat{f}_{+\infty}^\mu$ with the function f additional conditions are required.

Proposition 3. *Let f be a concave lower semicontinuous lower bounded function on the set $\mathfrak{S}(\mathcal{H})$, having some particular symmetry.*

A) *For each natural k the concave lower semicontinuous function $\hat{f}_k^\sigma = \hat{f}_k^\mu$ has the same symmetry and coincides with the function f on the set $\mathfrak{S}_k(\mathcal{H})$.*
The pointwise limit $\hat{f}_{+\infty}^\sigma = \hat{f}_{+\infty}^\mu$ of the monotonic sequence $\{\hat{f}_k^\sigma = \hat{f}_k^\mu\}$ coincides with the function f on the set $\mathfrak{S}(\mathcal{H})$.

B) *If the function f has continuous restriction to the set $\mathfrak{S}_k(\mathcal{H})$ for each natural k then the sequence $\{\hat{f}_k^\sigma = \hat{f}_k^\mu\}$ consists of concave continuous bounded functions on the set $\mathfrak{S}(\mathcal{H})$.*

Proof. By Proposition 1 $\hat{f}_k^\sigma = \hat{f}_k^\mu$ and this function is lower semicontinuous for each k. This implies $\hat{f}_{+\infty}^\sigma = \hat{f}_{+\infty}^\mu$ and lower semicontinuity of the last function. Since the function f is μ-concave by Proposition A-2 in the Appendix in [25], the first assertion of the proposition follows from the previous observations and Lemma 6 below.

The second assertion of the proposition follows from Corollary 1 since it is easy to see that continuity of the restrictions of the concave function f to the set $\mathfrak{S}_k(\mathcal{H})$ for all k implies boundedness of these restrictions. $\square$

Lemma 6. *A lower semicontinuous lower bounded concave function f on the set $\mathfrak{S}(\mathcal{H})$ is uniquely determined by its restriction to the set $\mathfrak{S}_\mathrm{f}(\mathcal{H})$ of finite rank states.*

[5] This means that the function f is invariant with respect to the particular family of symmetries of the set $\mathfrak{S}(\mathcal{H})$.

Proof. It is sufficient to consider the case of a nonnegative function f.

Let ρ_0 be an arbitrary state and let $\left\{\rho_n = (\mathrm{Tr} P_n \rho_0)^{-1} P_n \rho_0\right\}$ be the sequence of finite rank states converging to the state ρ_0, where $\{P_n\}$ is the sequence of finite rank spectral projectors of the state ρ_0 increasing to the identity operator $I_{\mathcal{H}}$.

For each n the inequality $\lambda_n \rho_n \leq \rho_0$ with $\lambda_n = \mathrm{Tr} P_n \rho_0$ 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. By concavity and nonnegativity of the function f we have $f(\rho_0) \geq \lambda_n f(\rho_n)$ for all n, which implies $\limsup_{n\to+\infty} f(\rho_n) \leq f(\rho_0)$. By lower semicontinuity of the function f we have $\lim_{n\to+\infty} f(\rho_n) = f(\rho_0)$. $\square$

Remark 2. The first assertion of Proposition 3 can be considered as a "constructive form" of Lemma 6 since it provides a constructive way of restoring a lower semicontinuous lower bounded concave function on the set $\mathfrak{S}(\mathcal{H})$ by means of its restriction to the set $\mathfrak{S}_{\mathrm{f}}(\mathcal{H})$.

Note that the above functions $\hat{f}^{\sigma}_{+\infty}$ and $\hat{f}^{\mu}_{+\infty}$ can be used in study of the following *construction problem*: for a given concave function defined on the convex set $\mathfrak{S}_{\mathrm{f}}(\mathcal{H})$ of finite rank states and having some particular analytical and symmetry properties to construct its concave extension to the set $\mathfrak{S}(\mathcal{H})$ of all states preserving these properties. Since in the proof of Proposition 3 the restriction of the function f to the set $\mathfrak{S}_{\mathrm{f}}(\mathcal{H})$ is only used, it shows that *for an arbitrary concave lower bounded function f on the set $\mathfrak{S}_{\mathrm{f}}(\mathcal{H})$ with some particular symmetry such that its restriction to the set $\mathfrak{S}_k(\mathcal{H})$ is lower semicontinuous for each k there exists a unique concave lower semicontinuous extension $\hat{f}^{\sigma}_{+\infty} = \hat{f}^{\mu}_{+\infty}$ to the set $\mathfrak{S}(\mathcal{H})$ with the same symmetry.* For example, if f is an entropy-type (t.i. nonnegative concave lower semicontinuous unitary invariant) function defined on the set of finite rank states, then $\hat{f}^{\sigma}_{+\infty} = \hat{f}^{\mu}_{+\infty}$ is its unique entropy-type extension to the set of all states. $\square$

The second assertion of Proposition 3 and the generalized Dini's lemma (in which the condition of continuity of functions of an increasing sequence is relaxed to their lower semicontinuity) imply the following continuity condition.

Corollary 2. *Let f be a concave lower semicontinuous lower bounded function on the set $\mathfrak{S}(\mathcal{H})$.*

A) *If the function f has continuous restriction to the set $\mathfrak{S}_k(\mathcal{H})$ for each k then uniform convergence of the sequence $\{\hat{f}^{\sigma}_k = \hat{f}^{\mu}_k\}$ on a particular subset $\mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$ implies continuity of the function f on this subset.*
B) *Continuity of the function f on a compact subset $\mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$ implies uniform convergence of the sequence $\{\hat{f}^{\sigma}_k = \hat{f}^{\mu}_k\}$ on this subset.*

We will use Corollary 2 in the next section to obtain continuity conditions for the von Neumann entropy.

5. The Approximation of the von Neumann Entropy and the Continuity Conditions

The von Neumann entropy $H(\rho) = -\mathrm{Tr}\rho \log \rho$ is a lower semicontinuous concave unitary invariant function on the set $\mathfrak{S}(\mathcal{H})$ of quantum states with the range $[0, +\infty]$, having continuous restriction to the set $\mathfrak{S}_k(\mathcal{H})$ for each k. By Proposition 3 the function H is a pointwise limit of the increasing sequence $\{H_k\}$ of nonnegative concave continuous bounded[6] unitary invariant functions on the set $\mathfrak{S}(\mathcal{H})$ defined as follows

$$H_k(\rho) = \sup_{\{\pi_i,\rho_i\}\in\mathcal{P}^{\mathrm{a}}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))} \sum_i \pi_i H(\rho_i) = \sup_{\mu\in\mathcal{P}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))} \int_{\mathfrak{S}_k(\mathcal{H})} H(\sigma)\mu(d\sigma).$$

(The first supremum is over all decompositions of the state ρ into countable convex combination of states of rank $\leq k$ while the second one is over all probability measures with the barycenter ρ supported by states of rank $\leq k$.)

For each k the function H_k may be called the *entropy approximator of order k* or briefly *k-approximator*. By construction the von Neumann entropy coincides with its k-approximator on the set $\mathfrak{S}_k(\mathcal{H})$ of all states of rank $\leq k$. For arbitrary state $\rho\in\mathfrak{S}(\mathcal{H})$ the difference $\Delta_k(\rho) = H(\rho) - H_k(\rho)$ between the von Neumann entropy and its k-approximator can be expressed as follows:

$$\Delta_k(\rho) = \inf_{\{\pi_i,\rho_i\}\in\mathcal{P}^{\mathrm{a}}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))} \sum_i \pi_i H(\rho_i\|\rho) = \inf_{\mu\in\mathcal{P}_{\{\rho\}}(\mathfrak{S}_k(\mathcal{H}))} \int_{\mathfrak{S}_k(\mathcal{H})} H(\sigma\|\rho)\mu(d\sigma),$$

where $H(\cdot\|\cdot)$ is the relative entropy [15,29] (the first equality follows from expression (4) below, the second one – from Proposition 1 in [11]). The possibility to express the value $\Delta_k(\rho)$ via the relative entropy is essentially used in what follows (see Lemma 8 below).

The representation of the von Neumann entropy as a limit of the increasing sequence $\{H_k\}$ of concave continuous bounded unitary invariant functions can be used for different purposes, in particular, for construction of the increasing sequence of continuous entanglement monotones providing approximation of the Entanglement of Formation (see Sect. 6 in [25]). By Corollary 2 this representation can be used for proving continuity of the von Neumann entropy on a subset of states by showing uniform convergence to zero of the sequence $\{\Delta_k\}$ on this subset. The last property of a subset of states, in what follows called the *uniform approximation property*, is considered in detail in the next subsection (in the extended context of subsets of the positive cone of trace-class operators).

5.1. The uniform approximation property. Since in many applications it is necessary to deal with the following extensions (cf.[13])

$$S(A) = -\mathrm{Tr}A\log A \quad \text{and} \quad H(A) = S(A) - \eta(\mathrm{Tr}A)$$

of the von Neumann entropy to the cone $\mathfrak{T}_+(\mathcal{H})$ of all positive trace-class operators (where $\eta(x) = -x\log x$), we will obtain the continuity conditions for the function $A\mapsto H(A)$ on this extended domain.

In what follows the function $A\mapsto H(A)$ on the cone $\mathfrak{T}_+(\mathcal{H})$ is called the *quantum entropy* while the function $\{x_i\}\mapsto H(\{x_i\}) = \sum_i \eta(x_i) - \eta\left(\sum_i x_i\right)$ on the positive cone of the space l_1, coinciding with the Shannon entropy on the set $\mathfrak{P}_{+\infty}$ of probability distributions, is called the *classical entropy*.

6 It is easy to see that the range of the function H_k coincides with $[0,\log k]$.

The von Neumann entropy has the important property expressed in the following inequality:

$$H\left(\sum_{i=1}^{n}\lambda_i\rho_i\right)\leq\sum_{i=1}^{n}\lambda_i H(\rho_i)+\sum_{i=1}^{n}\eta(\lambda_i),\tag{1}$$

valid for arbitrary set $\{\rho_i\}_{i=1}^n$ of states and probability distribution $\{\lambda_i\}_{i=1}^n$, where $n\leq+\infty$ (Proposition 6.2 in [15] and the simple approximation).

The definition and inequality (1) with $n=2$ imply the following properties of the quantum entropy:

$$H(\lambda A)=\lambda H(A),\tag{2}$$

$$H(A)+H(B-A)\leq H(B)\leq H(A)+H(B-A)+\mathrm{Tr}Bh_2\left(\frac{\mathrm{Tr}A}{\mathrm{Tr}B}\right),\tag{3}$$

where $A,B\in\mathfrak{T}_+(\mathcal{H}),\ A\leq B,\lambda\geq 0$ and $h_2(x)=\eta(x)+\eta(1-x)$.

Note that

$$S(A)-\sum_i\pi_i S(A_i)=\sum_i\pi_i H(A_i\,\|\,A)\tag{4}$$

for an arbitrary ensemble $\{\pi_i,A_i\}$ of operators in $\mathfrak{T}_+(\mathcal{H})$ with the average A, where $H(\cdot\|\cdot)$ is the (extended) relative entropy defined for arbitrary operators A and B in $\mathfrak{T}_+(\mathcal{H})$ as follows (cf.[13]):

$$H(A\,\|\,B)=\sum_i\langle i|\,(A\log A-A\log B+B-A)\,|i\rangle,$$

where $\{|i\rangle\}$ is the orthonormal basis of eigenvectors of A and it is assumed that $H(A\,\|\,B)=+\infty$ if $\mathrm{supp}A$ is not contained in $\mathrm{supp}B$. It is easy to verify that

$$H(\lambda A\,\|\,\lambda B)=\lambda H(A\,\|\,B),\quad\lambda\geq 0.\tag{5}$$

For given natural k consider the function

$$H_k(A)=\sup_{\{\pi_i,A_i\}\in\mathcal{P}^{\mathrm{a}}_{\{A\}}(\mathfrak{T}_+^k(\mathcal{H}))}\sum_i\pi_i H(A_i)$$

on the set $\mathfrak{T}_+(\mathcal{H})$ (the supremum is over all decompositions of the operator A into a countable convex combination of operators of rank $\leq k$). By using (2) it is easy to see that the restriction of the above function H_k to the set $\mathfrak{S}(\mathcal{H})$ coincides with the k-approximator of the von Neumann entropy defined in the first part of this section (so, we use the same notation) and that

$$H_k(\lambda A)=\lambda H_k(A),\quad A\in\mathfrak{T}_+(\mathcal{H}),\ \lambda\geq 0.$$

Thus we have

$$H_k(A)=\|A\|_1\hat{H}_k^{\sigma}(\|A\|_1^{-1}A)\leq\|A\|_1\log k,\quad A\in\mathfrak{T}_+(\mathcal{H}).\tag{6}$$

The contribution of the strong stability property of the set $\mathfrak{S}(\mathcal{H})$ to the below results is based on the following observation.

Lemma 7. *For arbitrary natural k the function $A \mapsto H_k(A)$ is continuous on the cone $\mathfrak{T}_+(\mathcal{H})$.*

Proof. By means of (6) the assertion of the lemma follows from Corollary 1 showing continuity of the function $\rho \mapsto \hat{H}_k^{\sigma}(\rho)$ on the set $\mathfrak{S}(\mathcal{H})$. □

For given natural k consider the function

$$\Delta_k(A) = \inf_{\{\pi_i, A_i\} \in \mathcal{P}^{\mathrm{a}}_{\{A\}}(\mathfrak{T}^k_+(\mathcal{H}))} \sum_i \pi_i H(A_i \| A) \tag{7}$$

on the set $\mathfrak{T}_+(\mathcal{H})$ (the infimum is over all decompositions of the operator A into a countable convex combination of operators of rank $\leq k$).

It follows from (5) that

$$\Delta_k(\lambda A) = \lambda \Delta_k(A), \quad A \in \mathfrak{T}_+(\mathcal{H}),\ \lambda \geq 0. \tag{8}$$

By Lemma 8 below the restriction of the function Δ_k defined in (7) to the set $\mathfrak{S}(\mathcal{H})$ coincides with the function $\Delta_k = H - H_k$ defined in the first part of this section (so, we use the same notation).

We will use the following properties of the function Δ_k.

Lemma 8. *For each natural k the following assertions hold:*

A) *For an arbitrary operator $A \in \mathfrak{T}_+(\mathcal{H})$ the infimum in definition (7) of the value $\Delta_k(A)$ can be taken over the subset of $\mathcal{P}^{\mathrm{a}}_{\{A\}}(\mathfrak{T}^k_+(\mathcal{H}))$ consisting of ensembles $\{\pi_i, A_i\}$ such that $\mathrm{Tr} A_i = \mathrm{Tr} A$ for all i and hence*

$$\Delta_k(A) = H(A) - H_k(A).$$

B) *The function $\mathfrak{T}_+(\mathcal{H}) \ni A \mapsto \Delta_k(A)$ is nonnegative lower semicontinuous unitary invariant and homogenous in the sense of (8). $\Delta_k^{-1}(0) = \mathfrak{T}^k_+(\mathcal{H})$. Continuity of this function on a subset $\mathcal{A} \subset \mathfrak{T}_+(\mathcal{H})$ means continuity of the quantum entropy on the subset $\mathcal{A}$.*

C) *The function $A \mapsto \Delta_k(A)$ is monotone with respect to the operator order:*

$$A \leq B \quad \Rightarrow \quad \Delta_k(A) \leq \Delta_k(B), \qquad \forall A,\ B \in \mathfrak{T}_+(\mathcal{H}).$$

D) *Let $\{\lambda_i(A)\}$ be the sequence of the eigenvalues of the operator $A \in \mathfrak{T}_+(\mathcal{H})$ arranged in nonincreasing order,[7] then*

$$\Delta_k(A) \leq \widetilde{\Delta}_k(A) \doteq H(\{\lambda_i^k(A)\}) = \sum_{i=1}^{+\infty} \eta(\lambda_i^k(A)) - \eta(\|A\|_1),$$

where the sequence $\{\lambda_i^k(A)\}$ is the k-order coarse-graining of the sequence $\{\lambda_i(A)\}$, t.i. $\lambda_i^k(A) = \lambda_{(i-1)k+1}(A) + \cdots + \lambda_{ik}(A)$ for all $i = 1, 2, \ldots$.

E) *For arbitrary operators A in $\mathfrak{T}_+(\mathcal{H})$ and C in $\mathfrak{B}(\mathcal{H})$ the following inequality holds:*

$$\Delta_k(CAC^*) \leq \|C\|^2 \Delta_k(A).$$

[7] It is possible to take the sequence $\{\lambda_i(A)\}$ in arbitrary order but the corresponding sequence $\{\lambda_i^k(A)\}$ is most close to the sequence $(\|A\|_1, 0, 0, \ldots)$ having zero entropy provided that the nonincreasing order is used. The relation between $\Delta_k(A)$ and $\widetilde{\Delta}_k(A)$ is considered in Remark 3 below.

F) *For an arbitrary operator A in $\mathfrak{T}_+(\mathcal{H})$ and an arbitrary sequence $\{P_n\}$ of projectors in $\mathfrak{B}(\mathcal{H})$ strongly converging to the identity operator $I_{\mathcal{H}}$ the following relation holds:*

$$\lim_{n\to+\infty} \Delta_k(P_n A P_n) = \Delta_k(A).$$

G) *For an arbitrary operator A in $\mathfrak{T}_+(\mathcal{H})$ and an arbitrary family $\{P_i\}_{i=1}^m$ of mutually orthogonal projectors in $\mathfrak{B}(\mathcal{H})$ ($m \leq +\infty$) the following inequality holds:*

$$\Delta_k(A) \geq \sum_{i=1}^{m} \Delta_k(P_i A P_i).$$

H) *For an arbitrary operator A in $\mathfrak{T}_+(\mathcal{H})$ and an arbitrary quantum operation $\Phi : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H})$ having the Kraus representation consisting of $\leq n$ summands the following inequality holds:*

$$\Delta_{nk}(\Phi(A)) \leq \Delta_k(A).$$

I) *For an arbitrary finite set $\{A_i\}_{i=1}^m$ of operators in $\mathfrak{T}_+(\mathcal{H})$ and corresponding set $\{k_i\}_{i=1}^m$ of natural numbers the following inequality holds:*

$$\Delta_{k_1+k_2+\cdots+k_m}\left(\sum_{i=1}^{m} A_i\right) \leq \sum_{i=1}^{m} \Delta_{k_i}(A_i).$$

J) *For an arbitrary countable set $\{A_i\}_{i=1}^{+\infty}$ of operators in $\mathfrak{T}_+(\mathcal{H})$, probability distribution $\{\lambda_i\}_{i=1}^{+\infty}$ and natural m the following inequality holds:*

$$\begin{aligned}\Delta_{mk}\left(\sum_{i=1}^{+\infty}\lambda_i A_i\right) &\leq \sum_{i=1}^{+\infty}\lambda_i \Delta_k(A_i) + \sum_{i=m}^{+\infty}\lambda_i H\left(A_i \,\Big\|\, \left(\sum_{i=m}^{+\infty}\lambda_i\right)^{-1}\sum_{i=m}^{+\infty}\lambda_i A_i\right)\\ &\leq \sum_{i=1}^{+\infty}\lambda_i \Delta_k(A_i)\\ &\quad + \sup_{i\geq m}\|A_i\|_1 H\left(\{\lambda_i\}_{i\geq m}\right).\end{aligned}$$

Proof. A) For arbitrary ensemble $\{\pi_i, A_i\}$ in $\mathcal{P}^{\mathrm{a}}_{\{A\}}(\mathfrak{T}^k_+(\mathcal{H}))$ one can consider ensemble $\{\lambda_i, B_i\}$ in $\mathcal{P}^{\mathrm{a}}_{\{A\}}(\mathfrak{T}^k_+(\mathcal{H}))$, where $\lambda_i = \pi_i\|A_i\|_1\|A\|_1^{-1}$ and $B_i = A_i\|A\|_1\|A_i\|_1^{-1}$, such that

$$\begin{aligned}\sum_i \lambda_i H(B_i\|A) &= \sum_i \pi_i H(A_i\|A)\\ &\quad -\left(\eta\left(\|A\|_1\right) - \sum_i \pi_i \eta\left(\|A_i\|_1\right)\right) \leq \sum_i \pi_i H(A_i\|A),\end{aligned}$$

where the last inequality follows from concavity of the function η, since $\sum_i \pi_i\|A_i\|_1 = \|A\|_1$. By (2) and (6) this implies $\Delta_k(A) = H(A) - H_k(A)$.

B) Lemma 7 and Assertion A imply the first and the third parts of this assertion. To prove the second one note that the inclusion $\mathfrak{T}^k_+(\mathcal{H}) \subseteq \Delta_k^{-1}(0)$ follows from the definition of the function Δ_k, while the converse inclusion is easily derived from the implication $\rho \in \mathfrak{S}(\mathcal{H}) \setminus \mathfrak{S}_k(\mathcal{H}) \Rightarrow H(\rho) > H_k(\rho)$, which follows from strict concavity of the von Neumann entropy and the last assertion of Proposition 2, implying attainability of the supremum in the second (continuous) expression in the definition of the function $H_k(\rho)$.

C) If $A \leq B$ then there exists contraction C such that $A = CBC^*$. Indeed, on the subspace $\mathrm{supp}B$ this contraction is constructed as the continuous extension to this subspace of the linear operator $A^{1/2}B^{-1/2}$ defined on the linear hull of the eigenvectors of the operator B corresponding to the positive eigenvalues, while on the subspace $\mathcal{H} \ominus \mathrm{supp}B$ it acts as the zero operator. Hence this assertion follows from Assertion H proved below.

D) Let $\{P_i^k\}_i$ be the sequence of spectral projectors of the operator A such that the projector P_i^k corresponds to the eigenvalues $\lambda_{(i-1)k+1}(A), \ldots, \lambda_{ik}(A)$. Then $\lambda_i^k(A) = \mathrm{Tr} P_i^k A$ for all i and the ensemble $\{\pi_i^k, (\pi_i^k)^{-1} P_i^k A\}$, where $\pi_i^k = \lambda_i^k(A)\|A\|_1^{-1}$, belongs to the set $\mathcal{P}^{\mathrm{a}}_{\{A\}}(\mathfrak{T}_+^k(\mathcal{H}))$. Hence

$$\Delta_k(A) \leq \sum_i \pi_i^k H((\pi_i^k)^{-1} P_i^k A \| A) = H(\{\lambda_i^k(A)\}).$$

E) By means of (8) this follows from Assertion H proved below.

F) By lower semicontinuity of the function Δ_k (Assertion B) this follows from Assertion E.

G) It is sufficient to prove that

$$\Delta_k(A) \geq \Delta_k(PAP) + \Delta_k(\bar{P}A\bar{P}),$$

where $\bar{P} = I_{\mathcal{H}} - P$, for arbitrary projector P. This inequality is easily proved by using the definition of the function Δ_k and the inequality

$$H(A\|B) \geq H(PAP\|PBP) + H(\bar{P}A\bar{P}\|\bar{P}B\bar{P})$$

valid for arbitrary operators A and B in $\mathfrak{T}_+(\mathcal{H})$ (Lemma 3 in [13]).

H) This follows from monotonicity of the relative entropy since for an arbitrary ensemble $\{\pi_i, A_i\}$ in $\mathcal{P}^{\mathrm{a}}_{\{A\}}(\mathfrak{T}_+^k(\mathcal{H}))$ the ensemble $\{\pi_i, \Phi(A_i)\}$ lies in $\mathcal{P}^{\mathrm{a}}_{\{\Phi(A)\}}(\mathfrak{T}_+^{nk}(\mathcal{H}))$.

I) By means of (8) it is sufficient to show that

$$\Delta_{k'+k''}(\gamma A + (1-\gamma)B) \leq \gamma \Delta_{k'}(A) + (1-\gamma)\Delta_{k''}(B) \tag{9}$$

for arbitrary operators A and B in $\mathfrak{T}_+(\mathcal{H})$ and $\gamma \in [0, 1]$. For given k' and k'' let $\{\pi_i, A_i\}_i$ and $\{\lambda_j, B_j\}_j$ be ensembles of operators of rank $\leq k'$ with the average A and of rank $\leq k''$ with the average B correspondingly. Then the ensemble $\{\pi_i\lambda_j, \gamma A_i + (1-\gamma)B_j\}_{i,j}$ has the average $\gamma A + (1-\gamma)B$ and consists of operators of rank $\leq k' + k''$. By joint convexity of the relative entropy we have

$$\begin{aligned}\Delta_{k'+k''}(\gamma A + (1-\gamma)B) &\leq \sum_{i,j} \pi_i\lambda_j H(\gamma A_i + (1-\gamma)B_j \| \gamma A + (1-\gamma)B)\\ &\leq \gamma \sum_i \pi_i H(A_i\|A) + (1-\gamma)\sum_j \lambda_j H(B_j\|B),\end{aligned}$$

which implies inequality (9).

J) The first inequality with $m = 1$ is easily derived from the definition of the function Δ_k by using Donald's identity

$$\sum_i \pi_i H(A_i \| B) = \sum_i \pi_i H(A_i \| A) + H(A \| B),$$

valid for arbitrary ensemble $\{\pi_i, A_i\}$ of positive trace-class operators with the average A and arbitrary trace-class operator B [15]. The case $m > 1$ is reduced to the case $m = 1$ by applying Assertion I (with (8)) to the sum $\sum_{i=1}^{m} A'_i$, where $A'_i = \lambda_i A_i$ for $i = \overline{1, m-1}$ and $A'_m = \sum_{i=m}^{+\infty} \lambda_i A_i$.

The second inequality follows from the estimation

$$\sum_i \pi_i H(A_i \| A) \leq \sup_i \|A_i\|_1 H(\{\pi_i\}),$$

valid for arbitrary ensemble $\{\pi_i, A_i\}$ of trace-class operators with the average A, which can be proved by using monotonicity of the relative entropy:

$$\sum_i \pi_i H(A_i \| A) = c \sum_i \pi_i H(\Phi(\rho_i) \| \Phi(\rho)) \leq c \sum_i \pi_i H(\rho_i \| \rho) = c H(\{\pi_i\}),$$

where $c = \sup_i \|A_i\|_1$, $\Phi(\cdot) = c^{-1} \sum_i \langle i| \cdot |i\rangle A_i$, $\rho_i = |i\rangle\langle i|$ and $\rho = \sum_i \pi_i \rho_i$. $\square$

Remark 3. It is easy to show that the upper bound $\widetilde{\Delta}_k(A)$ in Assertion D of Lemma 8 obtained by using the spectral decomposition of the operator A tends to zero if $H(A)$ is finite, which provides the additional proof of convergence of the sequence $\{H_k\}$ to the function H on the cone $\mathfrak{T}_+(\mathcal{H})$. Noncoincidence of the functions $\widetilde{\Delta}_k$ and Δ_k, t.i. existence of such operator A in $\mathfrak{T}_+(\mathcal{H})$ that $\Delta_k(A) < \widetilde{\Delta}_k(A)$, can be shown by the following example.

Let ρ be the chaotic state in a particular 3-D subspace $\mathcal{H}_0 \subset \mathcal{H}$. It is clear that $\widetilde{\Delta}_2(\rho) = \log 3 - \frac{2}{3}\log 2 \approx 0.64$ (we use the natural logarithm).

In the subspace $\mathcal{H}_0$ consider four unit vectors

$$|\varphi_1\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, |\varphi_2\rangle = \begin{bmatrix} -1/2 \\ \sqrt{3}/2 \\ 0 \end{bmatrix}, |\varphi_3\rangle = \begin{bmatrix} -1/2 \\ -\sqrt{3}/2 \\ 0 \end{bmatrix}, |\varphi_4\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

By direct calculation of eigenvalues one can show that the two rank states $\rho_1 = \frac{1}{2}|\varphi_1\rangle\langle\varphi_1| + \frac{1}{2}|\varphi_2\rangle\langle\varphi_2|$ and $\rho_2 = \frac{2}{5}|\varphi_3\rangle\langle\varphi_3| + \frac{3}{5}|\varphi_4\rangle\langle\varphi_4|$ have the entropies $H(\rho_1) \approx 0.57$ and $H(\rho_2) \approx 0.67$. Since $\frac{4}{9}\rho_1 + \frac{5}{9}\rho_2 = \rho$ we can conclude that $H_2(\rho) \geq \frac{4}{9}H(\rho_1) + \frac{5}{9}H(\rho_2) \approx 0.63$. Thus $\Delta_2(\rho) = H(\rho) - H_2(\rho) < \widetilde{\Delta}_2(\rho)$. $\square$

The following notion plays the central role in this paper.

Definition 1. *A subset $\mathcal{A}$ of $\mathfrak{T}_+(\mathcal{H})$ has the uniform approximation property (briefly the UA-property) if*

$$\lim_{k \to +\infty} \sup_{A \in \mathcal{A}} \Delta_k(A) = 0.$$

Importance of the UA-property is justified by its close relation to continuity of the quantum entropy considered in Theorem 2 in the next subsection. Usefulness of this relation is based on the following observation, showing the conservation of the UA-property under different set-operations.

Proposition 4. *Let $\mathcal{A}$ be a subset of $\mathfrak{T}_+(\mathcal{H})$ having the UA-property.*

A) *The UA-property holds for the closure* $\mathrm{cl}(\mathcal{A})$ *of the set* $\mathcal{A}$.

B) *For each* $\lambda > 0$ *the UA-property holds for the set*

$$M_\lambda(\mathcal{A}) = \{\lambda A \,|\, A \in \mathcal{A}\}.$$

C) *If* $\inf_{A\in\mathcal{A}} \|A\|_1 > 0$ *then the UA-property holds for the set*

$$E(\mathcal{A}) = \{\lambda A \mid A \in \mathcal{A}, \lambda \geq 0\} \cap \mathfrak{T}_1(\mathcal{H}).$$

D) *For each natural m the UA-property holds for the set*

$$\mathrm{co}_m(\mathcal{A}) = \left\{ \sum_{i=1}^{m} \pi_i A_i \,\middle|\, \{\pi_i\} \in \mathfrak{P}_m, \{A_i\} \subseteq \mathcal{A} \right\}.$$

If the set $\mathcal{A}$ is bounded then the UA-property holds for the set

$$\mathrm{co}_{\mathfrak{P}}(\mathcal{A}) = \left\{ \sum_{i=1}^{+\infty} \pi_i A_i \,\middle|\, \{\pi_i\} \in \mathfrak{P}, \{A_i\} \subseteq \mathcal{A} \right\},$$

where $\mathfrak{P}$ is a subset of $\mathfrak{P}_{+\infty}$ such that $\lim\limits_{m\to+\infty} \sup\limits_{\{\pi_i\}\in\mathfrak{P}} H(\{\pi_i\}_{i>m}) = 0.$

E) *The UA-property holds for the sets*

$$D(\mathcal{A}) = \{ B \in \mathfrak{T}_+(\mathcal{H}) \mid \exists A \in \mathcal{A} : B \leq A\}$$

and

$$\widetilde{D}(\mathcal{A}) = \{ B \in \mathfrak{T}_+(\mathcal{H}) \mid \exists A \in \mathcal{A} : B \lessdot A\},$$

where $B \lessdot A$ means that the sequence $\{\lambda_i(B)\}$ of eigenvalues of the operator B is majorized by the sequence $\{\lambda_i(A)\}$ of eigenvalues of the operator A in the sense $\lambda_i(B) \leq \lambda_i(A)$ for all i;

If the set $\mathcal{A}$ is compact and does not contain the null operator, then the UA-property holds for the set

$$\widehat{D}(\mathcal{A}) = \left\{ B \in \mathfrak{T}_1(\mathcal{H}) \mid \exists A \in \mathcal{A} : B\|B\|_1^{-1} \prec A\|A\|_1^{-1} \right\},$$

where $\rho \prec \sigma$ means that the state σ is more chaotic than the state ρ in the Uhlmann sense [1,30], *t.i. for the sequences $\{\lambda_i(\rho)\}$ and $\{\lambda_i(\sigma)\}$ of eigenvalues of the states ρ and σ arranged in nonicreasing order the inequality $\sum_{i=1}^{n} \lambda_i(\rho) \geq \sum_{i=1}^{n} \lambda_i(\sigma)$ holds for each natural n.*

F) *The UA-property holds for the sets*

$$Q_n(\mathcal{A}) = \{\Phi(A) \mid \Phi \in \mathfrak{F}^n_{\leq 1}(\mathcal{H}),\, A \in \mathcal{A}\}, \quad n \in \mathbb{N},$$

and

$$Q_{\mathfrak{F}}(\mathcal{A}) = \{\Phi(A) \mid \Phi \in \mathfrak{F},\, A \in \mathcal{A}\},$$

where $\mathfrak{F}$ *is a subset of* $\mathfrak{F}_{\leq 1}(\mathcal{H})$ *such that for the corresponding set* $\mathfrak{V}$ *of sequences* $\{V_j\}_{j=1}^{+\infty}$ *of Kraus operators the following two conditions hold:* [8]

1) *either* $\mathrm{Ran} V_j^* \perp \mathrm{Ran} V_{j'}^*$ *for all* $\{V_j\}_{j=1}^{+\infty} \in \mathfrak{V}$ *and all* $j \neq j'$ *exceeding some natural* n *or* $\lim_{m\to+\infty} \sup_{\{V_j\}\in\mathfrak{V}, A\in\mathcal{A}} \sum_{j\geq m} H\left(V_j A V_j^*\right) = 0$*;*
2)

$$\lim_{m\to+\infty} \sup_{\{V_j\}\in\mathfrak{V}, A\in\mathcal{A}} H\left(\{\mathrm{Tr} V_j A V_j^*\}_{j\geq m}\right) = 0.$$

Remark 4. In connection with Assertion D note that the UA-property of a set $\mathcal{A}$ does not imply the UA-property of its σ-convex hull $\sigma\text{-co}(\mathcal{A}) = \{\sum_{i=1}^{+\infty} \pi_i A_i | \{\pi_i\} \in \mathfrak{P}_{+\infty}, \{A_i\} \subseteq \mathcal{A}\}$ even if the set $\mathcal{A}$ is compact. As an example one can consider the converging sequence of pure states from Example 1 in the second part of [24], such that the von Neumann entropy is not continuous on the σ-convex hull of this sequence (since the UA-property implies continuity of the entropy by Lemma 7).

Note that the condition $\lim_{m\to+\infty} \sup_{\{\pi_i\}\in\mathfrak{P}} H(\{\pi_i\}_{i>m}) = 0$ means continuity of the classical entropy on the set $\mathfrak{P}$ provided that this set is compact.

Proof of Proposition 4. A) This follows from lower semicontinuity of the function Δ_k on the set $\mathfrak{T}_+(\mathcal{H})$ for each k (Lemma 8B).

B) This is an obvious corollary of (8).

C) This also follows from (8) since

$$\sup_{B\in E(\mathcal{A})} \{\lambda \mid B = \lambda A,\, A \in \mathcal{A}\} \leq \left(\inf_{A\in\mathcal{A}} \|A\|_1\right)^{-1}.$$

D) The first part follows from Lemma 8I and (8) implying

$$\Delta_{km}\left(\sum_{i=1}^{m} \pi_i A_i\right) \leq \sum_{i=1}^{m} \pi_i \Delta_k(A_i), \quad \forall \{\pi_i\}_{i=1}^m \in \mathfrak{P}_m.$$

The second part follows from Lemma 8J since for arbitrary k and m it implies

$$\begin{aligned}\Delta_{km}\left(\sum_{i=1}^{+\infty} \pi_i A_i\right) &\leq \sum_{i=1}^{+\infty} \pi_i \Delta_k(A_i) + \sup_{i\geq m} \|A_i\|_1 H\left(\{\pi_i\}_{i\geq m}\right) \\ &\leq \sup_{A\in\mathcal{A}} \Delta_k(A) + \sup_{A\in\mathcal{A}} \|A\|_1 H\left(\{\pi_i\}_{i\geq m}\right), \\ &\qquad \forall \{A_i\}_{i=1}^{+\infty} \subseteq \mathcal{A},\ \forall \{\pi_i\}_{i=1}^{+\infty} \in \mathfrak{P}_{+\infty}.\end{aligned}$$

8 The ways to show validity of these conditions are considered in the proof of Corollary 9 below.

E) The first part follows from Lemma 8C and unitary invariance of the function Δ_k. The second part follows from Lemma 8D and Lemma 9 below since

$$B\|B\|_1^{-1} \prec A\|A\|_1^{-1} \quad \Rightarrow \quad H(\{\lambda_i^k(B)\}) \leq \|B\|_1\|A\|_1^{-1} H(\{\lambda_i^k(A)\})$$

for each natural k by Shur concavity of the von Neumann entropy [30].

F) The first part follows from Lemma 8H.
To prove the second part note that Lemma 8J implies the inequality

$$\Delta_{km}\left(\sum_{j=1}^{+\infty} V_j A V_j^*\right) \leq \sum_{j=1}^{+\infty} \Delta_k(V_j A V_j^*) + H\left(\{\mathrm{Tr} V_j A V_j^*\}_{j\geq m}\right).$$

Thus it is sufficient to show that condition 1) implies

$$\lim_{k\to+\infty} \sup_{\{V_j\}\in\mathfrak{V}, A\in\mathcal{A}} \sum_{j=1}^{+\infty} \Delta_k(V_j A V_j^*) = 0. \tag{10}$$

If the first alternative in condition 1) holds then Assertions E and G of Lemma 8 provide the estimation

$$\begin{aligned}\sum_{j=1}^{+\infty} \Delta_k(V_j A V_j^*) &= \sum_{j=1}^{n} \Delta_k(V_j A V_j^*) + \sum_{j>n} \Delta_k(V_j A V_j^*)\\ &\leq n\Delta_k(A) + \sum_{j>n} \Delta_k(P_j A P_j) \leq (n+1)\Delta_k(A),\\ &\qquad \forall \{V_j\} \in \mathfrak{V},\end{aligned}$$

where P_j is the projector on the subspace $\mathrm{Ran}\, V_j^*$, which implies (10) by the UA-property of the set $\mathcal{A}$.

If the second alternative in condition 1) holds then the similar estimation, in which the term $\sum_{j>n} \Delta_k(V_j A V_j^*)$ is majorized by $\sum_{j>n} H(V_j A V_j^*)$ also implies (10) by the UA-property of the set $\mathcal{A}$. □

Lemma 9. *Let $\mathcal{A}$ be a compact subset of $\mathfrak{T}_+(\mathcal{H})$ having the UA-property. Then*

$$\lim_{k\to+\infty} \sup_{A\in\mathcal{A}} \widetilde{\Delta}_k(A) = 0$$

where $\widetilde{\Delta}_k$ is the upper bound for the function Δ_k defined in Lemma 8D.

Proof. By Lemma 7 the UA-property of the set $\mathcal{A}$ implies continuity of the function $A \mapsto H(A)$ on this set.

Let $\{P_i^k\}_i$ be the sequence of spectral projectors of the operator A defined in the proof of Assertion D of Lemma 8 and $\pi_i^k = \|A\|_1^{-1}\mathrm{Tr} P_i^k A$ for all i. By Lemma 4 in [13] the sequence of continuous functions $A \mapsto H(P_1^k A)$ monotonously converges to the function $A \mapsto H(A)$ as $k \to +\infty$. By Dini's lemma this sequence converges uniformly on the set $\mathcal{A}$. This implies the assertion of the lemma since

$$\widetilde{\Delta}_k(A) = \sum_i \pi_i^k H((\pi_i^k)^{-1} P_i^k A\|A) \leq H(A) - H(P_1^k A), \quad A \in \mathcal{A}.$$

□

By definition the UA-property of sets $\mathcal{A}$ and $\mathcal{B}$ implies the UA-property of their union $\mathcal{A} \cup \mathcal{B}$. By Lemma 8I and Proposition 4D we have the following observations.

Corollary 3. *Let $\mathcal{A}$ and $\mathcal{B}$ be subsets of $\mathfrak{T}_+(\mathcal{H})$ having the UA-property.*

A) *The UA-property holds for the set $\{A + B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$ (generally called the Minkowski sum of the sets $\mathcal{A}$ and $\mathcal{B}$);*
B) *The UA-property holds for the convex closure $\overline{\mathrm{co}}(\mathcal{A} \cup \mathcal{B})$ of the union of $\mathcal{A}$ and $\mathcal{B}$ provided these sets are convex.*

5.2. The continuity conditions. Lemmas 7 and 8, Dini's lemma and Proposition 4 imply the following theorem, containing the main results of this paper.

Theorem 2. A) *If a set $\mathcal{A} \subset \mathfrak{T}_+(\mathcal{H})$ has the UA-property then the quantum entropy is continuous on this set.*
B) *If the quantum entropy is continuous on a compact set $\mathcal{A} \subset \mathfrak{T}_+(\mathcal{H})$, then this set has the UA-property.*
C) *If a set $\mathcal{A} \subset \mathfrak{T}_+(\mathcal{H})$ has the UA-property then the quantum entropy is continuous on the set $\Lambda(\mathcal{A})$, where Λ is an arbitrary finite composition of the set-operations* cl, M_λ, E, co_m, $\mathrm{co}_{\mathfrak{P}}$, D, $\widetilde{D}$, $\widehat{D}$, Q_n, $Q_{\mathfrak{F}}$ *considered in Proposition 4 with arbitrary parameters $m, n \in \mathbb{N}$ and $\lambda > 0$ provided the sets $\mathfrak{P}$, $\mathfrak{F}$ and the arguments of E,* $\mathrm{co}_{\mathfrak{P}}$, $\widehat{D}$, $Q_{\mathfrak{F}}$ *satisfy the conditions mentioned in this proposition.*

Remark 5. As the simplest example showing importance of the compactness condition in the second assertion of Theorem 2, one can consider the set $\mathcal{A} = \{\lambda\rho \mid \lambda \in \mathbb{R}_+\}$, where ρ is an infinite rank state with finite entropy.

The following example shows that the second assertion of Theorem 2 can not be valid even for relatively compact convex sets of states.

Let $\{\rho_i\}_{i\geq 0}$ be a sequence of finite rank states in $\mathfrak{S}(\mathcal{H})$ such that ρ_0 is a pure state, $H(\rho_i) \geq 1$ for all $i > 0$, $\mathrm{supp}\rho_n \subset \mathcal{H} \ominus \left(\bigoplus_{i=0}^{n-1} \mathrm{supp}\rho_i\right)$ and $\sum_{i=1}^{+\infty} e^{-\lambda H(\rho_i)} < +\infty$ for all $\lambda > 0$. Let $\lambda_i = (H(\rho_i))^{-1}$ for each $i \in \mathbb{N}$. Consider the sequence of states

$$\sigma_i = (1 - \lambda_i)\rho_0 + \lambda_i \rho_i, \quad i \in \mathbb{N},$$

obviously converging to the state ρ_0.

In Appendix 7.2 it is proved that *the von Neumann entropy is continuous on the convex set $\mathcal{A} = \sigma\text{-co}\,(\{\sigma_i\}_{i\in\mathbb{N}}) = \left\{\sum_{i=1}^{+\infty} \pi_i\sigma_i \mid \{\pi_i\} \in \mathfrak{P}_{+\infty}\right\}$, but it is not continuous on the set* $\mathrm{cl}\,(\mathcal{A}) = \overline{\mathrm{co}}\,(\{\sigma_i\}_{i\in\mathbb{N}}) = \mathcal{A} \cup \{\rho_0\}$. By the first assertion of Theorem 2 and Proposition 4A the UA-property does not hold for the set $\mathcal{A}$. □

Show first that Theorem 2 makes it possible to re-derive the continuity conditions mentioned in the Introduction in the generalized forms.

Example 1. Let $\{h_i\}$ be a nondecreasing sequence of nonnegative numbers and $\mathfrak{P}_{\{h_i\},h}$ be the subset of $\mathfrak{P}_{+\infty}$ consisting of probability distributions $\{\pi_i\}$ satisfying the inequality $\sum_i h_i\pi_i \leq h$. By Lemma 11 in the Appendix the set $\mathfrak{P}_{\{h_i\},h}$ satisfies the condition in Proposition 4D if and only if $\mathrm{g}(\{h_i\}) = \inf\left\{\lambda > 0 \mid \sum_i e^{-\lambda h_i} < +\infty\right\} = 0$. By Theorem 2C the von Neumann entropy is continuous on the set $\mathrm{cl}(\mathrm{co}_{\mathfrak{P}_{\{h_i\},h}}(\mathfrak{S}_k(\mathcal{H})))$ for each k. This observation provides another proof[9] of the well known result stated

[9] The original proof of this result is based on lower semicontinuity of the function $\rho \mapsto H(\rho\|\sigma_\lambda)$, where $\sigma_\lambda = (\mathrm{Tr}e^{-\lambda H})^{-1}e^{-\lambda H}$, for all $\lambda > 0$ [15,29].

that the entropy is continuous on the set $\mathcal{K}_{H,h} = \{\rho \in \mathfrak{S}(\mathcal{H}) \mid \mathrm{Tr} H\rho \leq h\}$, where H is an $\mathfrak{H}$-operator such that $\mathrm{g}(H) = \inf\left\{\lambda > 0 \mid \mathrm{Tr} e^{-\lambda H} < +\infty\right\} = 0$, since by using the extremal properties of eigenvalues of a positive operator it is easy to see that the set $\mathrm{cl}(\mathrm{co}_{\mathfrak{P}_{\{h_i\},h}}(\mathfrak{S}_1(\mathcal{H})))$, where $\{h_i\}$ is the sequence of eigenvalues of the operator H, contains the set $\mathcal{K}_{H,h}$ (and all its unitary translations).

The von Neumann entropy is not continuous on $\mathrm{cl}(\mathrm{co}_{\mathfrak{P}_{\{h_i\},h}}(\mathfrak{S}_1(\mathcal{H}))$ if $\mathrm{g}(\{h_i\}) > 0$ since it is not continuous on the set $\mathcal{K}_{H,h}$ if $\mathrm{g}(H) > 0$ [24]. $\square$

Theorem 2 implies the following generalization of Simon's dominated convergence theorems [26].

Corollary 4 (Generalized Simon's convergence theorem). [10] *If the quantum entropy is continuous on a compact subset $\mathcal{A}$ of $\mathfrak{T}_+(\mathcal{H})$ then it is continuous on the sets $D(\mathcal{A})$ and $\widetilde{D}(\mathcal{A})$ defined in the first part of Assertion E of Proposition 4.*

This condition and Corollary 3 show that

$$\{H(A_n + B_n) \to H(A_0 + B_0)\} \Leftrightarrow \{H(A_n) \to H(A_0)\} \wedge \{H(B_n) \to H(B_0)\},$$

where $\{A_n\}$ and $\{B_n\}$ are sequences of positive trace class operators converging respectively to operators A_0 and B_0.

The above "dominated-type" continuity conditions can be enriched by the following one.

Corollary 5. *If the quantum entropy is continuous on a compact subset $\mathcal{A}$ of $\mathfrak{T}_+(\mathcal{H})$, which does not contain the null operator, then it is continuous on the set $\widehat{D}(\mathcal{A})$ defined in the second part of Assertion E of Proposition 4.*

By Corollary 5 and Theorem 13 in [30] to prove continuity of the von Neumann entropy on a set $\mathcal{A} \subset \mathfrak{S}(\mathcal{H})$ it suffices to show its continuity on the image of this set under the expectation $\rho \mapsto \sum_i P_i \rho P_i$ for a particular set $\{P_i\}$ of mutually orthogonal projectors such that $\sum_i P_i = I_{\mathcal{H}}$.

Corollary 5 and the infinite-dimensional generalization of Nielsen's theorem provide the following observation concerning the notion of entanglement of a state of a composite quantum system.

Example 2. Let $\mathcal{H}$ and $\mathcal{K}$ be separable Hilbert spaces. The entanglement $E(\omega)$ of a pure state ω in $\mathfrak{S}(\mathcal{H}\otimes\mathcal{K})$ is defined as the von Neumann entropy of its reduced states (cf.[2]):

$$E(\omega) = H(\mathrm{Tr}_{\mathcal{K}}\omega) = H(\mathrm{Tr}_{\mathcal{H}}\omega).$$

Let $\mathfrak{L}(\mathcal{H}, \mathcal{K})$ be the set of all LOCC-operations transforming the set $\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$ into itself. Corollary 5 and Lemma 2 in [17][11] imply the following assertion:

---

[10] In the original versions of these theorems the weaker topologies are used. Since the set $D(\mathcal{A})$ is compact (by the compactness criterion in Lemma 10 in the Appendix), the weak operator topology on this set coincides with the trace norm topology. The μ-convergence topology does not coincide with the trace norm topology on the set $\widetilde{D}(\mathcal{A})$, but by noting that the sequences of eigenvalues of the operators in $\widetilde{D}(\mathcal{A})$ form a compact subset of the space l_1 it is easy to see that μ-convergence of a sequence $\{A_n\} \subset \widetilde{D}(\mathcal{A})$ to an operator $A_0 \in \widetilde{D}(\mathcal{A})$ means trace norm convergence of the sequence $\{U_n A_n U_n^*\} \subset \widetilde{D}(\mathcal{A})$ to the operator A_0 for some set $\{U_n\}$ of unitaries.

[11] In [17] the majorization order is used, which is converse to the Uhlmann order "$\prec$" used in this paper.

If the function $\omega \mapsto E(\omega)$ *is continuous on a compact set* $\mathcal{C} \subset \mathrm{extr}\mathfrak{S}(\mathcal{H} \otimes \mathcal{K})$, *then it is continuous on the set*

$$\{\Lambda(\omega) \mid \omega \in \mathcal{C},\ \Lambda \in \mathfrak{L}(\mathcal{H}, \mathcal{K})\} \cap \mathrm{extr}\mathfrak{S}(\mathcal{H} \otimes \mathcal{K}).$$

This shows that for an arbitrary sequence $\{\omega_n\}$ of pure states in $\mathfrak{S}(\mathcal{H}\otimes\mathcal{K})$ converging to a state ω_0 and arbitrary set $\{\Lambda_n\}_{n\geq 0}$ of LOCC-operations such that the sequence $\{\Lambda_n(\omega_n)\}$ consists of pure states and converges to the state $\Lambda_0(\omega_0)$, the following implication holds:

$$\lim_{n\to+\infty} E(\omega_n) = E(\omega_0) \quad \Rightarrow \quad \lim_{n\to+\infty} E(\Lambda_n(\omega_n)) = E(\Lambda_0(\omega_0)). \ \square$$

By Corollary 10 in the Appendix for arbitrary closed set $\mathcal{A} \subset \mathfrak{T}_+(\mathcal{H})$ (not necessarily compact) and arbitrary natural m the set $\mathrm{co}_m(\mathcal{A})$ defined in Assertion D of Proposition 4 is closed. Theorem 2 implies the following result.

Corollary 6. A) *If the quantum entropy is continuous and bounded on a closed bounded set* $\mathcal{A} \subset \mathfrak{T}_+(\mathcal{H})$ *then it is continuous on the set* $\mathrm{co}_m(\mathcal{A})$ *for arbitrary natural* m.

B) *If the quantum entropy is continuous on a compact set* $\mathcal{A} \subset \mathfrak{T}_+(\mathcal{H})$, *then it is continuous on the set* $\mathrm{cl}(\mathrm{co}_{\mathfrak{P}}(\mathcal{A}))$ *for arbitrary subset* $\mathfrak{P}$ *of* $\mathfrak{P}_{+\infty}$ *such that* $\lim\limits_{m\to+\infty} \sup\limits_{\{\pi_i\}\in\mathfrak{P}} H(\{\pi_i\}_{i>m}) = 0.$

By Remark 4 the set $\mathrm{cl}(\mathrm{co}_{\mathfrak{P}}(\mathcal{A}))$ in the second assertion of this corollary can not be replaced by the σ-convex hull σ-$\mathrm{co}(\mathcal{A})$ of the set $\mathcal{A}$.

Proof. A) Let $\{A_n\} \subset \mathrm{co}_m(\mathcal{A})$ be a sequence converging to an operator $A_0 \in \mathrm{co}_m(\mathcal{A})$. Suppose

$$\lim_{n\to+\infty} H(A_n) > H(A_0). \tag{11}$$

By the construction of the set $\mathrm{co}_m(\mathcal{A})$ for each n there exists an ensemble $\{\pi_i^n, A_i^n\}_{i=1}^m$ of operators in $\mathcal{A}$ such that $A_n = \sum_{i=1}^m \pi_i^n A_i^n$. By using Proposition 5 in the Appendix and boundedness of the set $\mathcal{A}$ we may consider (by replacing the sequence $\{A_n\}$ by some subsequence) that there exists an ensemble $\{\pi_i^0, A_i^0\}_{i=1}^m$ of operators in $\mathcal{A}$ such that $\lim_{n\to+\infty} \pi_i^n = \pi_i^0$, $\lim_{n\to+\infty} \pi_i^n A_i^n = \pi_i^0 A_i^0$ for each $i = \overline{1, m}$ and $A_0 = \sum_{i=1}^m \pi_i^0 A_i^0$.

Since the entropy is continuous and bounded on the set $\mathcal{A}$ we have $\lim_{n\to+\infty} H(\pi_i^n A_i^n) = H(\pi_i^0 A_i^0)$ for each $i = \overline{1, m}$. By the part "$\Leftarrow$" of the remark after Corollary 4 this implies a contradiction to (11).

B) This directly follows from Theorem 2. $\square$

If $\mathcal{A}$ is a union of $m < +\infty$ closed convex sets then Corollary 10 in the Appendix implies $\mathrm{co}_m(\mathcal{A}) = \overline{\mathrm{co}}(\mathcal{A})$, so we obtain from Corollary 6 the following result.

Corollary 7. *If the quantum entropy is continuous on each set from a finite collection* $\{\mathcal{A}_i\}_{i=1}^m$ *of convex closed bounded subsets of* $\mathfrak{T}_+(\mathcal{H})$ *then it is continuous on the convex closure* $\overline{\mathrm{co}}\left(\bigcup_{i=1}^m \mathcal{A}_i\right)$ *of this collection.*

Remark 6. The condition of closedness of the *all* sets from the collection $\{\mathcal{A}_i\}_{i=1}^m$ in Corollary 7 is essential. The simple example showing this can be constructed as follows. Let $\mathcal{A}_1 = \{\rho_0\}$ and $\mathcal{A}_2 = \sigma$-$\mathrm{co}\,(\{\sigma_i\}_{i\in\mathbb{N}})$, where the state ρ_0 and the sequence $\{\sigma_i\}_{i\in\mathbb{N}}$ are taken from the example in Remark 5. As shown in this example the entropy is continuous on the convex bounded sets $\mathcal{A}_1$ and $\mathcal{A}_2$ but it is not continuous on the convex set $\mathcal{A}_1 \cup \mathcal{A}_2$. $\square$

Theorem 2 also implies the following continuity condition.

Corollary 8. *Let $\{\mathcal{A}_i\}_{i=1}^n$ be a finite collection of subsets of $\mathfrak{T}_+(\mathcal{H})$ having the UA-property (for example, compact subsets on which the quantum entropy is continuous). Then for arbitrary natural m the quantum entropy is continuous on the set*

$$\mathrm{cl}\left(\left\{\sum_{i=1}^{n}\sum_{j=1}^{m} V_{ij}A_iV_{ij}^* \mid A_i \in \mathcal{A}_i,\, V_{ij} \in \mathfrak{B}(\mathcal{H}),\, \|V_{ij}\| \leq 1\right\}\right).$$

The following observation can be used in the study of quantum channels and in the theory of quantum measurements (see Example 3 below).

Corollary 9. *Let $\mathfrak{V}_{=1}$ be the set of all sequences $\{V_i\}_{i=1}^{+\infty} \subset \mathfrak{B}(\mathcal{H})$ such that $\sum_{i=1}^{+\infty} V_i^*V_i = I_{\mathcal{H}}$ endowed with the Cartesian product of the strong* operator topology (the topology of coordinate-wise strong* operator convergence).*[12] *Let $\mathcal{A}$ be a subset of $\mathfrak{T}_+(\mathcal{H})$ on which the quantum entropy is continuous.*

1) *The function $(\{V_i\}, A) \mapsto \sum_{i=1}^{+\infty} H\left(V_iAV_i^*\right)$ is continuous on $\mathfrak{V}_{=1} \times \mathcal{A}$.*
2) *If $\mathfrak{V}_0$ is a subset of $\mathfrak{V}_{=1}$ such that the function*

$$(\{V_i\}, A) \mapsto H\left(\left\{\mathrm{Tr}V_iAV_i^*\right\}_{i=1}^{+\infty}\right) \tag{12}$$

is continuous on $\mathfrak{V}_0 \times \mathcal{A}$ then the function $(\{V_i\}, A) \mapsto H\left(\sum_{i=1}^{+\infty} V_iAV_i^*\right)$ is continuous on $\mathfrak{V}_0 \times \mathcal{A}$.

Proof. We can consider that the sets $\mathcal{A}$ and $\mathfrak{V}_0$ are compact.

1) It follows from Corollary 8 that the function $F_m((\{V_i\}, A)) = H(C_mAC_m)$, where $C_m = \sqrt{\sum_{i=1}^m V_i^*V_i}$ and $m \in \mathbb{N}$, is continuous on $\mathfrak{V}_{=1} \times \mathcal{A}$. Since $C_m^2 \leq C_{m+1}^2$ for all m the sequence $\{F_m\}$ is nondecreasing. By noting that convergence of a sequence $\{A_n\} \subset \mathfrak{T}_+(\mathcal{H})$ to an operator $A_0 \in \mathfrak{T}_+(\mathcal{H})$ follows from its convergence in the weak operator topology provided that $\lim_n \mathrm{Tr}A_n = \mathrm{Tr}A_0$ (Theorem 1 in [6]) and by using Corollary 8 we conclude that $\lim_{m\to+\infty} F_m((\{V_i\}, A)) = H(A)$.

By the Groenevold-Lindblad-Ozawa inequality (see [16]) we have

$$\sum_{i>m} H(V_iAV_i^*) \leq H(A) - F_m((\{V_i\}, A)).$$

Hence continuity of the function $A \mapsto H(A)$ and Dini's lemma show that $\lim_{m\to+\infty} \sup_{\{V_i\}\in\mathfrak{V}_c, A\in\mathcal{A}} \sum_{i>m} H(V_iAV_i^*) = 0$ for an arbitrary compact subset $\mathfrak{V}_c$ of $\mathfrak{V}_{=1}$. This and continuity of the function $(\{V_i\}, A) \mapsto \sum_{i=1}^m H\left(V_iAV_i^*\right)$ for each m (provided by Corollary 8) imply the first assertion of the corollary.

2) By the above observation the second alternative in condition 1) in Proposition 4F holds for the sets $\mathfrak{V}_0$ and $\mathcal{A}$. Since condition 2) in this proposition follows from continuity of function (12) by Dini's lemma, the set $\left\{\sum_{i=1}^{+\infty} V_iAV_i^* \mid \{V_i\} \in \mathfrak{V}_0, A \in \mathcal{A}\right\}$ has the UA-property. By Theorem 2 this implies the second assertion of the corollary. □

[12] The strong* operator topology on $\mathfrak{B}(\mathcal{H})$ is defined by the family of seminorms $A \mapsto \|A|\varphi\rangle\| + \|A^*|\varphi\rangle\|$, $|\varphi\rangle \in \mathcal{H}$ [3]. By using more complicated analysis it is possible to replace this topology here by the strong operator topology.

Example 3. Let $\mathfrak{M}_m(\mathcal{H})$ be the set of all quantum measurements with $m \leq +\infty$ outcomes on the quantum system associated with the Hilbert space $\mathcal{H}$. Each measurement $\mathcal{M}$ *in* $\mathfrak{M}_m(\mathcal{H})$ is described by a set $\{V_i\}_{i=1}^m$ of operators in $\mathfrak{B}(\mathcal{H})$ such that $\sum_{i=1}^m V_i^* V_i = I_{\mathcal{H}}$ and its action on an arbitrary a priori state $\rho \in \mathfrak{S}(\mathcal{H})$ results in the posteriori ensemble $\{\pi_i(\mathcal{M}, \rho), \rho_i(\mathcal{M}, \rho)\}_{i=1}^m$, where $\rho_i(\mathcal{M}, \rho) = (\mathrm{Tr} V_i \rho V_i^*)^{-1} V_i \rho V_i^*$ is the posteriori state corresponding to the i^{th} outcome and $\pi_i(\mathcal{M}, \rho) = \mathrm{Tr} V_i \rho V_i^*$ is the probability of this outcome (if $\mathrm{Tr} V_i \rho V_i^* = 0$ then the posteriori state $\rho_i(\mathcal{M}, \rho)$ is not defined) [10]. The mean posteriori state $\bar{\rho}(\mathcal{M}, \rho) = \sum_{i=1}^m \pi_i(\mathcal{M}, \rho)\rho_i(\mathcal{M}, \rho) = \sum_{i=1}^m V_i \rho V_i^*$ corresponds to the nonselective measurement.

We will consider that a sequence $\{\mathcal{M}_n\} \subset \mathfrak{M}_m(\mathcal{H})$ converges to a measurement $\mathcal{M}_0 \in \mathfrak{M}_m(\mathcal{H})$ if $\lim_{n\to+\infty} V_i^n = V_i^0$ for all $i = \overline{1, m}$ in the strong* operator topology, where $\{V_i^n\}_{i=1}^m$ is the set of operators describing the measurement $\mathcal{M}_n$.

Let $\mathcal{A}$ be a subset of $\mathfrak{S}(\mathcal{H})$ on which the von Neumann entropy is continuous. Corollaries 8 and 9 imply the following assertions:

- the von Neumann entropy of the posteriori state $H(\rho_i(\mathcal{M}, \rho))$ is continuous on the subset of $\mathfrak{M}_m(\mathcal{H}) \times \mathcal{A}$, on which $\rho_i(\mathcal{M}, \rho)$ is defined, $i = \overline{1, m}$;
- the mean entropy of posteriori states $\sum_{i=1}^m \pi_i(\mathcal{M}, \rho) H(\rho_i(\mathcal{M}, \rho))$ is continuous on $\mathfrak{M}_m(\mathcal{H}) \times \mathcal{A}$;
- if $\mathfrak{M}$ is a subset of $\mathfrak{M}_m(\mathcal{H})$ such that the Shannon entropy of the outcome's probability distribution $H\left(\{\pi_i(\mathcal{M}, \rho)\}_{i=1}^m\right)$ is continuous on $\mathfrak{M} \times \mathcal{A}$ then the von Neumann entropy of the mean posteriori state $H(\bar{\rho}(\mathcal{M}, \rho))$ is continuous on $\mathfrak{M} \times \mathcal{A}$.

If $m < +\infty$, then the function $H(\bar{\rho}(\mathcal{M}, \rho))$ is continuous on $\mathfrak{M}_m(\mathcal{H}) \times \mathcal{A}$. $\square$

Remark 7. The continuity conditions considered in this subsection are formulated for subsets of $\mathfrak{T}_+(\mathcal{H})$. They can be reformulated for subsets of $\mathfrak{S}(\mathcal{H})$ by using the following obvious observation: *If the quantum entropy is continuous on a subset* $\mathcal{A}$ *of* $\mathfrak{T}_+(\mathcal{H})$ *such that* $\inf_{A\in\mathcal{A}} \|A\|_1 > 0$ *then the von Neumann entropy is continuous on the subset* $\left\{A\|A\|_1^{-1} \mid A \in \mathcal{A}\right\}$ *of* $\mathfrak{S}(\mathcal{H})$.

6. Conclusion

The method of proving continuity of the von Neumann entropy proposed in this paper is essentially based on the strong stability property (stated in Theorem 1) and on the μ-compactness (described before Lemma 1) of the set of quantum states, revealing the special relations between the topology and the convex structure of this set. Of course, it does not mean that validity of the continuity conditions obtained by this method depends on validity of these abstract properties and that these conditions can not be proved by other methods. For example, the assertion of Corollary 7 for sets of quantum states can be shown by noting that continuity of the entropy on any closed convex set of states implies compactness of this set (this follows from Lemma 2 in [25] and Corollary 7 in [24]) and by applying spectral finite dimensional approximation based on using inequality (1) and Dini's lemma, but the proposed method provides a simpler and in a sense more natural way of doing this.

The special approximation of concave lower semicontinuous functions considered in this paper, in particular, the approximation of the von Neumann entropy used in proving its continuity, seems to be interesting for other applications.

7. Appendix

7.1. One property of the positive cone of trace-class operators. The positive cone $\mathfrak{T}_+(\mathcal{H})$ has the following important property.

Proposition 5. *Let* $\{\{\pi_i^n, A_i^n\}_{i=1}^m\}_n$ *be a sequence of ensembles consisting of* $m < +\infty$ *operators in* $\mathfrak{T}_+(\mathcal{H})$ *such that the sequence* $\{\sum_{i=1}^m \pi_i^n A_i^n\}_n$ *of their averages converges to an operator* A_0. *There exists a subsequence* $\{\{\pi_i^{n_k}, A_i^{n_k}\}_{i=1}^m\}_k$ *converging to a particular ensemble* $\{\pi_i^0, A_i^0\}_{i=1}^m$ *with the average* A_0 *in the following sense*:

$$\lim_{k\to+\infty} \pi_i^{n_k} = \pi_i^0 \quad and \quad \pi_i^0 > 0 \;\Rightarrow\; \lim_{k\to+\infty} A_i^{n_k} = A_i^0, \quad i = \overline{1,m}.$$

Note that this proposition does not assert that $A_i^0 \neq A_j^0$ for all $i \neq j$.

Proof. We may assume that the sequence $\{A_n = \sum_{i=1}^m \pi_i^n A_i^n\}_n$ belongs to the set $\mathfrak{T}_1(\mathcal{H})$ of positive operators with trace ≤ 1.

Let $B_i^n = \pi_i^n A_i^n$ be an operator in $\mathfrak{T}_1(\mathcal{H})$ such that $B_i^n \leq A_n$ for all $n > 0$ and $i = \overline{1,m}$. Since the set $\{A_n\}_{n\geq 0}$ is compact, the compactness criterion for subsets of $\mathfrak{T}_1(\mathcal{H})$ in Lemma 10 below implies relative compactness of the sequence $\{B_i^n\}_{n>0}$ for each $i = \overline{1,m}$. Hence we can find an increasing sequence $\{n_k\}$ of natural numbers such that there exist

$$\lim_{k\to+\infty} \pi_i^{n_k} = \pi_i^0 \quad \text{and} \quad \lim_{k\to+\infty} B_i^{n_k} = B_i^0 \quad \forall i = \overline{1,m},$$

where $\{\pi_i^0\}_{i=1}^m$ is a probability distribution and $\{B_i^0\}_{i=1}^m$ is a collection of operators in $\mathfrak{T}_1(\mathcal{H})$. Since $\sum_{i=1}^m B_i^{n_k} = A_{n_k}$ for all k we have $\sum_{i=1}^m B_i^0 = A_0$.

Let $A_i^0 = (\pi_i^0)^{-1} B_i^0$ if $\pi_i^0 \neq 0$ and $A_i^0 = 0$ otherwise. The subsequence $\{\{\pi_i^{n_k}, A_i^{n_k}\}_{i=1}^m\}_k$ and the ensemble $\{\pi_i^0, A_i^0\}_{i=1}^m$ have the required properties. $\square$

Corollary 10. *For an arbitrary closed subset* $\mathcal{A}$ *of* $\mathfrak{T}_+(\mathcal{H})$ *and an arbitrary natural* m *the set* $\mathrm{co}_m(\mathcal{A}) = \left\{\sum_{i=1}^m \pi_i A_i \,|\{\pi_i\} \in \mathfrak{P}_m, \{A_i\} \subset \mathcal{A}\right\}$ *is closed.*

The following compactness criterion for subsets of $\mathfrak{T}_1(\mathcal{H})$ can be derived from the compactness criterion for subsets of $\mathfrak{S}(\mathcal{H})$, presented in [11, the Appendix] (by considering the set $\{A + (1 - \mathrm{Tr}A)\rho_0 \,|\, A \in \mathcal{A}\} \subset \mathfrak{S}(\mathcal{H})$ for a given set $\mathcal{A} \subset \mathfrak{T}_1(\mathcal{H})$, where ρ_0 is a fixed pure state).

Lemma 10. *A closed subset* $\mathcal{A}$ *of* $\mathfrak{T}_1(\mathcal{H})$ *is compact if and only if for arbitrary* $\varepsilon > 0$ *there exists a finite rank projector* P_ε *in* $\mathfrak{B}(\mathcal{H})$ *such that* $\mathrm{Tr}(I_{\mathcal{H}} - P_\varepsilon)A < \varepsilon$ *for all* $A \in \mathcal{A}$.

7.2. The proofs of the auxiliary results. On maximality of any measure supported by pure states. Maximality of a measure μ in $\mathcal{P}(\mathfrak{S}_1(\mathcal{H}))$ with respect to the Choquet ordering follows from coincidence of this ordering with the dilation ordering [7], but it can be easily proved in this special case as follows. Suppose $\nu \succ \mu$. Since for an arbitrary concave continuous bounded function f on the $\mathfrak{S}(\mathcal{H})$ stability of the set $\mathfrak{S}(\mathcal{H})$ implies continuity of its convex hull $\mathrm{co} f$ (see [20, Theorem 1]) we have

$$\mu(f) \geq \nu(f) \;\text{ and }\; \mu(\mathrm{co} f) \leq \nu(\mathrm{co} f), \;\text{ where }\; \mu(f) = \int_{\mathfrak{S}(\mathcal{H})} f(\sigma)\mu(d\sigma).$$

By noting that $f \geq \mathrm{co} f$ and that these functions coincide on the support of the measure μ we conclude that $\mu(f) = \nu(f)$ and hence $\mu = \nu$.

The proof of Lemma 2. It is easy to see that lower semicontinuity and lower boundedness of the function f imply lower semicontinuity of the functional

$$\mathcal{P}(\mathcal{A}) \ni \mu \mapsto F(\mu) = \int_{\mathcal{A}} f(\sigma)\mu(d\sigma). \tag{13}$$

A) Convexity of the function $\check{f}_{\mathcal{A}}$ follows from its definition. By lower semicontinuity of the functional (13) and compactness of the set $\mathcal{P}_{\{\rho\}}(\mathcal{A})$ (provided by μ-compactness of the set $\mathfrak{S}(\mathcal{H})$) the infimum in the definition of the value $\check{f}_{\mathcal{A}}(\rho)$ for each ρ in $\overline{\mathrm{co}}(\mathcal{A})$ is achieved at a particular measure in $\mathcal{P}_{\{\rho\}}(\mathcal{A})$.

Suppose the function $\check{f}_{\mathcal{A}}$ is not lower semicontinuous. Then there exists a sequence $\{\rho_n\} \subset \overline{\mathrm{co}}(\mathcal{A})$ converging to a state $\rho_0 \in \overline{\mathrm{co}}(\mathcal{A})$ such that

$$\lim_{n\to+\infty} \check{f}_{\mathcal{A}}(\rho_n) < \check{f}_{\mathcal{A}}(\rho_0). \tag{14}$$

As proved before for each $n = 1, 2, \ldots$ there exists a measure μ_n in $\mathcal{P}_{\{\rho_n\}}(\mathcal{A})$ such that $\check{f}_{\mathcal{A}}(\rho_n) = F(\mu_n)$. μ-compactness of the set $\mathfrak{S}(\mathcal{H})$ implies existence of a subsequence $\{\mu_{n_k}\}$ converging to a particular measure μ_0. By continuity of the map $\mu \mapsto \mathbf{b}(\mu)$ the measure μ_0 belongs to the set $\mathcal{P}_{\{\rho_0\}}(\mathcal{A})$. Lower semicontinuity of functional (13) implies

$$\check{f}_{\mathcal{A}}(\rho_0) \leq F(\mu_0) \leq \liminf_{k\to+\infty} F(\mu_{n_k}) = \lim_{k\to+\infty} \check{f}_{\mathcal{A}}(\rho_{n_k}),$$

contradicting (14).

B) Concavity of the function $\hat{f}_{\mathcal{A}}$ follows from its definition. Suppose the function $\hat{f}_{\mathcal{A}}$ is not lower semicontinuous. Then there exists a sequence $\{\rho_n\} \subset \overline{\mathrm{co}}(\mathcal{A})$ converging to a state $\rho_0 \in \overline{\mathrm{co}}(\mathcal{A})$ such that

$$\lim_{n\to+\infty} \hat{f}_{\mathcal{A}}(\rho_n) < \hat{f}_{\mathcal{A}}(\rho_0). \tag{15}$$

For $\varepsilon > 0$ let μ_0^{ε} be such a measure in $\mathcal{P}_{\{\rho_0\}}(\mathcal{A})$ that $\hat{f}_{\mathcal{A}}(\rho_0) < F(\mu_0^{\varepsilon}) + \varepsilon$. By openness of the map $\mathcal{P}(\mathcal{A}) \ni \mu \mapsto \mathbf{b}(\mu)$ there exists a subsequence $\{\rho_{n_k}\}$ and a sequence $\{\mu_k\} \subset \mathcal{P}(\mathcal{A})$ converging to the measure μ_0^{ε} such that $\mathbf{b}(\mu_k) = \rho_{n_k}$ for each k. Lower semicontinuity of functional (13) implies

$$\hat{f}_{\mathcal{A}}(\rho_0) \leq F(\mu_0^{\varepsilon}) + \varepsilon \leq \liminf_{k\to+\infty} F(\mu_k) + \varepsilon \leq \lim_{k\to+\infty} \hat{f}_{\mathcal{A}}(\rho_{n_k}) + \varepsilon,$$

contradicting (15) (since ε is arbitrary). $\square$

The proof of the assertion in Remark 5. For an arbitrary state ρ in $\mathcal{A} = \sigma\text{-co}(\{\sigma_i\})$ there exists a probability distribution $\{\pi_i\} \in \mathfrak{P}_{+\infty}$ such that $\rho = \sum_{i=1}^{+\infty} \pi_i\sigma_i$. This distribution is unique since $P_i\rho = \pi_i\lambda_i\rho_i$ for each i, where P_i is the projector on the subspace $\mathrm{supp}\rho_i$.

The one-to-one correspondence $\mathfrak{P}_{+\infty} \ni \{\pi_i\} \leftrightarrow \sum_i \pi_i\sigma_i \in \mathcal{A}$ is continuous in both directions (t.i. it is a homeomorphism). Indeed, continuity of the map "$\rightarrow$" is obvious while continuity of the map "$\leftarrow$" can be proved by using the above set $\{P_i\}$ of projectors and by noting that pointwise convergence of a sequence of probability distributions to a *probability distribution* implies its convergence in the norm of total variation.

Thus to prove continuity of the von Neumann entropy on the set $\mathcal{A}$ it is sufficient to show continuity of the function $\mathfrak{P}_{+\infty} \ni \{\pi_i\} \mapsto H\left(\sum_i \pi_i\sigma_i\right)$.

By the construction of the sequence $\{\sigma_i\}$ we have

$$\begin{aligned} H\left(\sum_{i=1}^{+\infty}\pi_i\sigma_i\right) &= H\left(\left(\sum_{i=1}^{+\infty}\pi_i(1-\lambda_i)\right)\rho_0 \oplus \bigoplus_{i=1}^{+\infty}\pi_i\lambda_i\rho_i\right) \\ &= \sum_{i=1}^{+\infty}\pi_i(1-\lambda_i)H(\rho_0) + \sum_{i=1}^{+\infty}\pi_i\lambda_i H(\rho_i) \\ &\quad +\eta\left(\sum_{i=1}^{+\infty}\pi_i(1-\lambda_i)\right) + \sum_{i=1}^{+\infty}\eta(\pi_i\lambda_i) \\ &= 1 + \eta\left(\sum_{i=1}^{+\infty}\pi_i(1-\lambda_i)\right) \\ &\quad + \sum_{i=1}^{+\infty}\pi_i\lambda_i(-\log\pi_i) + \sum_{i=1}^{+\infty}\pi_i\lambda_i(-\log\lambda_i). \end{aligned}$$

By using properties of the function $x \mapsto \eta(x)$ and Lemma 11 below it is easy to show continuity of the all terms in the right side of the above expression as functions of $\{\pi_i\}$.

Discontinuity of the von Neumann entropy on the set $\mathrm{cl}(\mathcal{A}) = \mathcal{A} \cup \{\rho_0\}$ follows from the inequality $H(\sigma_i) \geq \lambda_i H(\rho_i) = 1$, $i > 0$, since $H(\rho_0) = 0$.

Lemma 11. *Let $\{h_j\}_{j=1}^{+\infty}$ be a nondecreasing sequence of positive numbers such that* $\mathrm{g}\left(\{h_j\}\right) = \inf\left\{\lambda > 0 \mid \sum_{j=1}^{+\infty} e^{-\lambda h_j} < +\infty\right\} < +\infty$. *Then*

$$\lim_{m\to+\infty}\sup_{\{x_j\}\in\mathcal{B}_1}\sum_{j\geq m}\eta(x_j)h_j^{-1} = \mathrm{g}\left(\{h_j\}\right),$$

where $\mathcal{B}_1$ is the positive part of the unit ball of the Banach space l_1.

Proof. We will prove first that

$$\lambda_* \leq \sup_{\{x_j\}\in\mathcal{B}_1}\sum_{j=1}^{+\infty}\eta(x_j)h_j^{-1} \leq \lambda_* + h_1^{-1}, \tag{16}$$

where λ_* is either the unique solution of the equation $\sum_{j=1}^{+\infty} e^{-\lambda h_j} = e$ if it exists or $\mathrm{g}\left(\{h_j\}\right)$ otherwise (if $\sum_{j=1}^{+\infty} e^{-\mathrm{g}(\{h_j\})h_j} < e$).

By using the Lagrange method it is easy to show that the function $\{x_j\}_{j=1}^n \mapsto \sum_{j=1}^n \eta(x_j)h_j^{-1}$ attains its maximum on the positive part $\mathcal{B}_1^n$ of the unit ball of $\mathbb{R}^n$ at the vector $\{e^{-\lambda_n h_j - 1}\}_{j=1}^n$, where λ_n is the unique solution of the equation $\sum_{j=1}^n e^{-\lambda h_j} = e$, and hence

$$\lambda_n \leq \sup_{\{x_j\}\in\mathcal{B}_1^n}\sum_{j=1}^{n}\eta(x_j)h_j^{-1} = \lambda_n + \sum_{j=1}^{n} e^{-\lambda_n h_j - 1}h_j^{-1} \leq \lambda_n + h_1^{-1}.$$

It is easy to see that the increasing sequence $\{\lambda_n\}$ converges to λ_*, so by noting that $\{x_j\} \mapsto \sum_{j=1}^{+\infty} \eta(x_j)h_j^{-1}$ is a lower semicontinuous function and by passing to the limit in the above expression we obtain (16).

The assertion of the lemma can be derived from (16) applied to the sequence $\{h_{j+m}\}_{j=1}^{+\infty}$, since if the solution of the equation $\sum_{j=1}^{+\infty} e^{-\lambda h_{j+m}} = e$ exists for all m it tends to $\mathrm{g}\left(\{h_j\}\right)$ as $m \to +\infty$. $\square$

Acknowledgements. I am grateful to A.S.Holevo for the help and useful discussion. I am also grateful to I.Nechita for the information about the generalization of Nielsen's theorem and to the referees for the useful remarks. This work is partially supported by the program "Mathematical control theory" of Russian Academy of Sciences, by the federal target program "Scientific and pedagogical staff of innovative Russia" (program 1.2.1, contract P 938), by the analytical departmental target program "Development of scientific potential of the higher school 2009-2010" (project 2.1.1/500) and by RFBR grants 09-01-00424-a and 10-01-00139-a. I am grateful to the organizers of the workshop *Thematic Program on Mathematics in Quantum Information* at the Fields Institute, where some work improving the paper was done.

References

1. Alberti, P.M., Uhlmann, A.: *Stochasticity and Partial Order. Doubly Stochastic Maps and Unitary Mixing*. Berlin: VEB Deutscher Verlag Wiss., 1981
2. Bennett, C.H., DiVincenzo, D.P., Smolin, J.A., Wootters, W.K.: Mixed state entanglement and quantum error correction. Phys. Rev. A **54**, 3824–3851 (1996)
3. Bratteli, O., Robinson, D.W.: *Operators Algebras and Quantum Statistical Mechanics*. vol. **I**. New York-Heidelberg-Berlin: Springer Verlag, 1979
4. O'Brien, R.: On the openness of the barycentre map. Math. Ann. **223**(3), 207–212 (1976)
5. Clausing, A., Papadopoulou, S.: Stable convex sets and extremal operators. Math. Ann. **231**, 193–203 (1978)
6. Dell'Antonio, G.F.: On the limits of sequences of normal states. Commun. Pure Appl. Math. **20**, 413–430 (1967)
7. Edgar, G.A.: *On the Radon-Nikodim Property and Martingale Convergence*. Lecture Notes in Mathematics **645**, Berlin-Heidelberg: Springer, 1978, pp. 62–76
8. Grzaslewicz, R.: Extreme continuous function property. Acta. Math. Hungar. **74**, 93–99 (1997)
9. Holevo, A.S.: On quasi-equivalence of locally normal states. Theor. Math. Phys. **13**(2), 184–199 (1972)
10. Holevo, A.S.: *Statistical Structure of Quantum Theory*. Berlin-Heidelberg-New York: Springer-Verlag, 2001
11. Holevo, A.S., Shirokov, M.E.: Continuous ensembles and the χ-capacity of infinite dimensional channels. Theor. Prob. and Appl. **50**(1), 8698 (2005)
12. Joffe, A.D., Tikhomirov, W.M.: *Theory of Extremum Problems*. New York: Academic Press, 1979
13. Lindblad, G.: Expectation and entropy inequalities for finite quantum systems. Commun. Math. Phys. **39**(2), 111–119 (1974)
14. Nielsen, M.A., Chuang, I.L.: *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press, 2000
15. Ohya, M., Petz, D.: *Quantum Entropy and Its Use*. Texts and Monographs in Physics. Berlin: Springer-Verlag, 1993
16. Ozawa, M.: On information gain by quantum measurement of continuous observable. J. Math. Phys. **27**, 759–763 (1986)
17. Owari, M., Braunstein, S.L., Nemoto, K., Murao, M.: ε-convertibility of entangled states and extension of Schmidt rank in infinite-dimensional systems. Quant. Inf. Comp. **8**(1–2), 0030–0052 (2008)
18. Papadopoulou, S.: On the geometry of stable compact convex sets. Math. Ann. **229**, 193–200 (1977)
19. Parthasarathy, K.: *Probability Measures on Metric Spaces*. New York-London: Academic Press, 1967
20. Protasov, V.Yu., Shirokov, M.E.: Generalized compactness in linear spaces and its applications. Sbornik: Math. **200**(5), 697–722 (2009)
21. Plenio, M.B., Virmani, S.: An introduction to entanglement measures. Quant. Inf. Comp. **7**, 1 (2007)
22. Rockafellar, R.: *Convex Analysis*. Princeton, NJ: Princeton Univ. Press, 1970
23. Shirokov, M.E.: The Holevo capacity of infinite dimensional channels and the additivity problem. Commun. Math. Phys. **262**, 137–159 (2006)
24. Shirokov, M.E.: Entropic characteristics of subsets of states-I. Izvestiya: Math. **70**(6), 1265–1292 (2006)

25. Shirokov M.E.: *On properties of the space of quantum states and their application to construction of entanglement monotones*. Izvestiya: Math., to appear, available at http://arxiv.org/abs/0804.1515v2[math-ph], 2009
26. Simon, B.: *Convergence Theorem for Entropy*. Appendix in E.H.Lieb, M.B.Ruskai, Proof of the strong suadditivity of quantum mechanical entropy. J. Math. Phys. **14**, 1938 (1973)
27. Uhlmann, A.: Entropy and optimal decomposition of states relative to a maximal commutative subalgebra. Open Syst. Inf. Dyn. **5**(3), 209–228 (1998)
28. Wagner, D.H.: *Survey of Measurable Selection Theorems*. Lecture Notes in Mathematics **794**, Berlin: Springer, 1980, pp. 176–219
29. Wehrl, A.: General properties of entropy. Rev. Mod. Phys. **50**, 221–250 (1978)
30. Wehrl, A.: How chaotic is a state of a quantum system. Rep. Math. Phys. **6**, 15–28 (1974)

Communicated by M.B. Ruskai