THEORY PROBAB. APPL.
Vol. 50, No. 1, pp. 86–98

CONTINUOUS ENSEMBLES AND THE CAPACITY OF INFINITE-DIMENSIONAL QUANTUM CHANNELS*

A. S. HOLEVO† AND M. E. SHIROKOV†

(*Translated by A. S. Holevo*)

Abstract. This paper is devoted to the study of χ-capacity, closely related to the classical capacity of infinite-dimensional quantum channels. For such channels generalized ensembles are defined as probability measures on the set of all quantum states. We establish the compactness of the set of generalized ensembles with averages in an arbitrary compact subset of states. This result enables us to obtain a sufficient condition for the existence of the optimal generalized ensemble for an infinite-dimensional channel with input constraint. This condition is shown to be fulfilled for Bosonic Gaussian channels with constrained mean energy. In the case of convex constraints, a characterization of the optimal generalized ensemble extending the "maximal distance property"is obtained.

Key words. quantum channel, χ-capacity, generalized ensemble

DOI. 10.1137/S0040585X97981470

1. Introduction. This paper is devoted to the systematic study of the classical capacity (more precisely, a closely related quantity — the χ-capacity [6]) of infinite-dimensional quantum channels, following [8], [10], [17]. While major attention in quantum information theory up to now has been paid to finite-dimensional systems, there is an important and interesting class of Gaussian channels (see, e.g., [9], [4], [16]) which act in infinite-dimensional Hilbert space. Although many problems of Gaussian bosonic systems with a finite number of modes can be solved with finite-dimensional matrix techniques, a general underlying Hilbert space operator analysis is indispensable.

Moreover, it was observed recently [17] that Shor's famous proof of the global equivalence of different forms of the additivity conjecture is related to the weird discontinuity of the χ-capacity in the infinite-dimensional case. All this calls for a mathematically rigorous treatment involving specific results from the operator theory in Hilbert space and measure theory.

There are two important features essential for channels in infinite dimensions. One is the necessity of the input constraints (such as the mean energy constraint for Gaussian channels) to prevent infinite capacities (although considering input constraints was recently shown to be quite useful also in the study of the additivity conjecture for channels in finite dimensions [10]). Another is the natural appearance of infinite, and, in general, "continuous" state ensembles understood as probability measures on the set of all quantum states. By using compactness criteria from probability theory and operator theory we can show that the set of all generalized ensembles with the average in a compact set of states is itself a compact subset of the set of all probability measures. With this in hand we give a sufficient condition for the existence of an optimal generalized ensemble for a constrained quantum channel. This condition can be

*Received by the editors September 21, 2004. This work was partially supported by the grants INTAS 00-738 and NSh 1758.2003.1. The authors acknowledge support from the program "Quantum Information Science" at the Newton Institute, Cambridge, where this paper was completed.
http://www.siam.org/journals/tvp/50-1/98147.html
†Steklov Mathematical Institute RAS, Gubkin St. 8, 119991 Moscow, Russia (holevo@mi.ras.ru; msh@mi.ras.ru).

verified in particular in the case of bosonic Gaussian channels with constrained mean energy. In the case of convex constraints we give a characterization of the optimal generalized ensemble extending the "maximal distance property" (see [15], [10]).

2. Preliminaries. We give below for reference some results from noncommutative probability theory (see details in [7] and [14]).

Let $\mathcal{H}$ be a separable Hilbert space, $\mathfrak{B}(\mathcal{H})$ the algebra of all bounded operators in $\mathcal{H}$, and $\mathfrak{T}(\mathcal{H})$ the Banach space of all trace-class operators with the trace norm $\|\cdot\|_1$. *State* is a positive trace class operator ρ in $\mathcal{H}$ with unit trace: $\rho \geqq 0$, $\operatorname{Tr}\rho = 1$. The algebra $\mathfrak{B}(\mathcal{H})$ is called the algebra of observables of a quantum system; then a state determines the expectation functional $A \mapsto \operatorname{Tr}\rho A$, $A \in \mathfrak{B}(\mathcal{H})$. The set of all states $\mathfrak{S}(\mathcal{H})$ is a convex closed subset of $\mathfrak{T}(\mathcal{H})$ which is a complete separable metric space with the metric defined by the norm.

In what follows we shall use the fact that convergence of a sequence of states to a *state* in the weak operator topology is equivalent to the trace norm convergence [1]. Note also the following characterization of compact subsets of states (a noncommutative analogue of Prokhorov's theorem): A closed subset $\mathcal{K}$ of states in $\mathfrak{S}(\mathcal{H})$ is compact if and only if for any $\varepsilon > 0$ there is a finite-dimensional projector P such that $\operatorname{Tr}\rho P \geqq 1 - \varepsilon$ for all $\rho \in \mathcal{K}$. A proof of this result is given in Appendix A.

A finite set π_i, ρ_i of the states ρ_i with respective probabilities π_i is called an *ensemble*; the state $\overline{\rho} = \sum_i \pi_i\rho_i$ is called an average of the ensemble.

Definition. *We call an arbitrary Borel probability measure π on $\mathfrak{S}(\mathcal{H})$ a* generalized ensemble. *The* average[1] *of the generalized ensemble π is defined by the Pettis integral*

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

Using the result of [1] it is possible to show that the integral also exists in Bochner's sense (see [5]) with respect to the trace norm. The conventional ensembles correspond to measures with finite support.

Denote by $\mathcal{P}$ the convex set of all probability measures on $\mathfrak{S}(\mathcal{H})$ equipped with the topology of weak convergence [2]. The mapping $\pi \mapsto \bar{\rho}(\pi)$ is continuous in this topology. In fact the weak convergence of a sequence $\{\pi_n\} \subset \mathcal{P}$ to $\pi \in \mathcal{P}$ implies convergence of the sequence of states $\{\bar{\rho}(\pi_n)\}$ to the state $\bar{\rho}(\pi)$ in the weak operator topology, and, hence, by the result in [1], in the trace norm.

Lemma 1. *The subset of measures with finite support is dense in the set of all measures with given average $\bar{\rho}$.*

A proof of this statement is given in Appendix B.

In what follows, log denotes the function on $[0, +\infty)$, which coincides with the usual logarithm on $(0, +\infty)$ and vanishes at zero. If A is a positive finite rank operator, then the *entropy* is defined as

$$H(A) = \operatorname{Tr} A(I \log \operatorname{Tr} A - \log A), \tag{1}$$

where I is the unit operator in $\mathcal{H}$. In particular, the entropy of a state ρ (von Neumann entropy) is equal to

$$H(\rho) = -\operatorname{Tr}\rho\log\rho.$$

[1] Also called a barycenter of the measure π.

If A and B are two such operators, then the *relative entropy* is defined as

$$H(A\,\|\,B) = \operatorname{Tr}(A\log A - A\log B + B - A) \tag{2}$$

provided $\operatorname{ran} A \subseteq \operatorname{ran} B$, and $H(A\,\|\,B) = +\infty$ otherwise (throughout this paper, ran denotes the closure of the range of an operator).

These definitions can be extended to arbitrary positive trace class operators A and B with the help of the following lemma [11].

LEMMA 2. *Let $\{P_n\}$ be an arbitrary sequence of finite-dimensional projectors monotonously increasing to the unit operator I in the strong operator topology. The sequences $\{H(P_nAP_n)\}$ and $\{H(P_nAP_n\,\|\,P_nBP_n)\}$ are monotonously increasing and their limits* (*finite or infinite*) *do not depend on the choice of the sequence $\{P_n\}$.*

We can thus define the entropy and the relative entropy as

$$H(A) = \lim_{n\to+\infty} H(P_nAP_n), \qquad H(A\,\|\,B) = \lim_{n\to+\infty} H(P_nAP_n\,\|\,P_nBP_n).$$

As is well known, the properties of the entropy for infinite- and finite-dimensional Hilbert spaces differ quite substantially: In the latter case the entropy is a bounded continuous function on $\mathfrak{S}(\mathcal{H})$, while in the former it is discontinuous (lower semicontinuous) at every point, and infinite "almost everywhere" in the sense that the set of states with finite entropy is a first category subset of $\mathfrak{S}(\mathcal{H})$ [19].

3. The χ-capacity of constrained channels. Let $\mathcal{H},\mathcal{H}'$ be a pair of separable Hilbert spaces which we shall call correspondingly the input and the output space. A channel Φ is a linear positive trace-preserving map from $\mathfrak{T}(\mathcal{H})$ to $\mathfrak{T}(\mathcal{H}')$ such that the dual map $\Phi^*\colon \mathfrak{B}(\mathcal{H}') \mapsto \mathfrak{B}(\mathcal{H})$ (which exists since Φ is bounded) is completely positive [7]. In particular, a channel maps (input) states in $\mathcal{H}$ to (output) states in $\mathcal{H}'$.

Let $\mathcal{A}$ be an arbitrary subset of $\mathfrak{S}(\mathcal{H})$. Consider the constraint on an input ensemble $\{\pi_i,\rho_i\}$, defined by the requirement $\bar\rho \in \mathcal{A}$. The channel Φ with this constraint is called the $\mathcal{A}$-*constrained* channel. We define the χ-*capacity* of the $\mathcal{A}$-*constrained channel* Φ as

$$\overline{C}(\Phi;\mathcal{A}) = \sup_{\bar\rho\in\mathcal{A}} \chi_\Phi\bigl(\{\pi_i,\rho_i\}\bigr), \tag{3}$$

where

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

Throughout this paper we shall consider the constraint sets $\mathcal{A}$ such that

$$\overline{C}(\Phi;\,\mathcal{A}) < +\infty. \tag{5}$$

The subset of $\mathcal{P}$ consisting of all measures π with the average state $\bar\rho(\pi)$ in $\mathcal{A} \subseteq \mathfrak{S}(\mathcal{H})$ will be denoted by $\mathcal{P}_\mathcal{A}$.

Lemma 2 implies, in particular, that the nonnegative function

$$\rho \mapsto H\bigl(\Phi(\rho)\,\|\,\Phi\bigl(\bar\rho(\pi)\bigr)\bigr)$$

is measurable on $\mathfrak{S}(\mathcal{H})$. Hence the functional

$$\chi_\Phi(\pi) = \int_{\mathfrak{S}(\mathcal{H})} H\bigl(\Phi(\rho)\,\|\,\Phi\bigl(\bar\rho(\pi)\bigr)\bigr)\,\pi(d\rho)$$

is well defined on the set $\mathcal{P}$ (with the range $[0;+\infty]$).

Proposition 1. *The functional $\chi_\Phi(\pi)$ is lower semicontinuous on the set $\mathcal{P}$. If $H(\Phi(\bar\rho(\pi)))<\infty$, then*

$$\chi_\Phi(\pi)=H\big(\Phi\big(\bar\rho(\pi)\big)\big)-\int_{\mathfrak{S}(\mathcal{H})}H\big(\Phi(\rho)\big)\,\pi(d\rho). \tag{6}$$

Proof. Let $\{P_n\}$ be an arbitrary sequence of finite-dimensional projectors monotonously increasing to the unit operator I. We show first that the functionals

$$\chi_\Phi^n(\pi)=\int_{\mathfrak{S}(\mathcal{H})}H\big(P_n\Phi(\rho)\,P_n\,\|\,P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)\,\pi(d\rho)$$

are continuous on the set $\mathcal{P}$.

We have

$$\operatorname{ran}\big(P_n\Phi(\rho)\,P_n\big)\subseteq\operatorname{ran}\big(P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)$$

for π-almost all ρ. Indeed, closure of the range is the orthogonal complement to the null subspace of a Hermitian operator, and for the null subspaces the opposite inclusion holds obviously. It follows that

$$\begin{aligned}&H\big(P_n\Phi(\rho)\,P_n\,\|\,P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)\\&\quad=\operatorname{Tr}\Big(P_n\Phi(\rho)\,P_n\log\big(P_n\Phi(\rho)\,P_n\big)-P_n\Phi(\rho)\,P_n\log\big(P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)\\&\qquad\quad+P_n\Phi\big(\bar\rho(\pi)\big)\,P_n-P_n\Phi(\rho)\,P_n\Big)\end{aligned}$$

for π-almost all ρ. By using (1) we have

$$\begin{aligned}\chi_\Phi^n(\pi)=&-\int_{\mathfrak{S}(\mathcal{H})}H\big(P_n\Phi(\rho)\,P_n\big)\,\pi(d\rho)+\int_{\mathfrak{S}(\mathcal{H})}\operatorname{Tr}\big(P_n\Phi(\rho)\big)\,\log\operatorname{Tr}\big(P_n\Phi(\rho)\big)\,\pi(d\rho)\\&-\int_{\mathfrak{S}(\mathcal{H})}\operatorname{Tr}\big(P_n\Phi(\rho)\,P_n\big)\,\log\big(P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)\,\pi\,(d\rho)\\&+\int_{\mathfrak{S}(\mathcal{H})}\operatorname{Tr}\big(P_n\Phi\big(\bar\rho(\pi)\big)\big)\,\pi(d\rho)-\int_{\mathfrak{S}(\mathcal{H})}\operatorname{Tr}\big(P_n\Phi(\rho)\big)\,\pi(d\rho).\end{aligned}$$

It is easy to see that the two last terms cancel, while the central term can be transformed in the following way:

$$\begin{aligned}&-\int_{\mathfrak{S}(\mathcal{H})}\operatorname{Tr}\big(P_n\Phi(\rho)\,P_n\big)\,\log\big(P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)\,\pi(d\rho)\\&\quad=-\operatorname{Tr}\int_{\mathfrak{S}(\mathcal{H})}\big(P_n\Phi(\rho)\,P_n\big)\,\log\big(P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)\,\pi(d\rho)\\&\quad=H\big(P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)-\operatorname{Tr}\big(P_n\Phi\big(\bar\rho(\pi)\big)\big)\,\log\operatorname{Tr}\big(P_n\Phi\big(\bar\rho(\pi)\big)\big).\end{aligned}$$

Hence

$$\begin{aligned}\chi_\Phi^n(\pi)=&\,H\big(P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big)-\operatorname{Tr}\big(P_n\Phi\big(\bar\rho(\pi)\big)\big)\,\log\operatorname{Tr}\big(P_n\Phi\big(\bar\rho(\pi)\big)\big)\\&-\int_{\mathfrak{S}(\mathcal{H})}H\big(P_n\Phi(\rho)\,P_n\big)\,\pi(d\rho)+\int_{\mathfrak{S}(\mathcal{H})}\operatorname{Tr}\big(P_n\Phi(\rho)\big)\,\log\operatorname{Tr}\big(P_n\Phi(\rho)\big)\,\pi(d\rho).\end{aligned}\tag{7}$$

The continuity and boundedness of the quantum entropy in the finite-dimensional case and similar properties of the function $\rho\mapsto\operatorname{Tr}(P_n\Phi(\rho))\log\operatorname{Tr}(P_n\Phi(\rho))$ imply continuity of the functionals $\chi_\Phi^n(\pi)$ for all n.

By the monotonous convergence theorem the sequence of functionals $\chi^n_\Phi(\pi)$ is monotonously increasing and pointwise converges to $\chi_\Phi(\pi)$. Hence the functional $\chi_\Phi(\pi)$ is lower semicontinuous as an upper bound of a family of continuous functionals.

To prove (6) note that Lemma 2 implies

$$\lim_{n\to+\infty} H\big(P_n\Phi\big(\bar\rho(\pi)\big)\,P_n\big) = H\big(\Phi\big(\bar\rho(\pi)\big)\big)$$

and

$$\lim_{n\to+\infty}\int_{\mathfrak{S}(\mathcal{H})} H\big(P_n\Phi(\rho)\,P_n\big)\,\pi(d\rho) = \int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho)\big)\,\pi(d\rho)$$

due to the monotonous convergence theorem. For every ρ the sequence $\{\operatorname{Tr}(P_n\Phi(\rho))\}$ is in $[0,1]$ and converges to 1 and therefore, $\lim_{n\to+\infty}\operatorname{Tr}(P_n(\rho))\log\operatorname{Tr}(P_n(\rho)) = 0$; in particular the second term in (7) tends to 0. Since $|x\log x| < 1$ for all $x\in(0,1]$, the last term in (7) also tends to 0 by the dominated convergence theorem. So, passing to the limit $n\to\infty$ in (7) gives (6). Proposition 1 is proved.

COROLLARY 1. *The χ-capacity of $\mathcal{A}$-constrained channel Φ can be defined by*

$$\overline{C}(\Phi;\,\mathcal{A}) = \sup_{\pi\in\mathcal{P}_\mathcal{A}}\chi_\Phi(\pi).$$

Proof. Definition (3) is a similar expression in which the supremum is over all measures in $\mathcal{P}_\mathcal{A}$ with finite support. By Lemma 1 we can approximate arbitrary measure π in $\mathcal{P}_\mathcal{A}$ by a sequence $\{\pi_n\}$ of measures in $\mathcal{P}_\mathcal{A}$ with finite supports. By Proposition 1 $\liminf_{n\to+\infty}\chi_\Phi(\pi_n) \geqq \chi_\Phi(\pi)$. It follows that the supremum over all measures in $\mathcal{P}_\mathcal{A}$ coincides with the supremum over all measures in $\mathcal{P}_\mathcal{A}$ with finite support. Corollary 1 is proved.

4. Compact constraints. It will be convenient to use the following terminology: an unbounded positive operator H in $\mathcal{H}$ with discrete spectrum of finite multiplicity will be called an $\mathfrak{H}$-*operator*. Let Q_n be the spectral projector of H corresponding to the lowest n eigenvalues. Following [8] we shall denote

$$\operatorname{Tr}\rho H = \lim_{n\to\infty}\operatorname{Tr}\rho Q_n H, \tag{8}$$

where the sequence on the right side is monotonously nondecreasing. It was shown in [8] that

$$\mathcal{K} = \{\rho\colon\ \operatorname{Tr}\rho H \leqq h\} \tag{9}$$

is a compact subset of $\mathfrak{S}(\mathcal{H})$ for the *arbitrary* $\mathfrak{H}$-operator H.

LEMMA 3. *Let $\mathcal{A}$ be a compact subset of $\mathfrak{S}(\mathcal{H})$. Then there exist an $\mathfrak{H}$-operator H and a positive number h such that $\operatorname{Tr}\rho H \leqq h$ for all $\rho\in\mathcal{A}$.*

Proof. By the compactness criterion (see Appendix A) for any natural n there exists a finite rank projector P_n such that $\operatorname{Tr}\rho P_n \geqq 1-n^{-3}$ for all ρ in $\mathcal{A}$. Without loss of generality we may assume that $\bigvee_{k=1}^{+\infty}P_k(\mathcal{H}) = \mathcal{H}$, where $\bigvee$ denotes a closed linear span of the subspaces. Let $\widehat{P}_n$ be the projector on the finite-dimensional subspace $\bigvee_{k=1}^{n}P_k(\mathcal{H})$. Thus $H = \sum_{n=1}^{+\infty} n(\widehat{P}_{n+1}-\widehat{P}_n)$ is an $\mathfrak{H}$-operator satisfying

$$\operatorname{Tr}\rho H = \sum_{n=1}^{+\infty} n\operatorname{Tr}\rho(\widehat{P}_{n+1}-\widehat{P}_n) \leqq \sum_{n=1}^{+\infty} n\operatorname{Tr}\rho(I_\mathcal{H}-\widehat{P}_n) \leqq \sum_{n=1}^{+\infty} n^{-2} = h$$

for arbitrary state ρ in the set $\mathcal{A}$. Lemma 3 is proved.

This lemma can be used to establish compactness of some subsets of states. Consider, for example, the set $\mathcal{C}(\rho,\sigma)$ of all states ω in the tensor product of Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ having fixed partial traces $\operatorname{Tr}_{\mathcal{K}}\omega=\rho$ and $\operatorname{Tr}_{\mathcal{H}}\omega=\sigma$. By Lemma 3 there exist $\mathfrak{H}$-operators A and B in the spaces $\mathcal{H}$ and $\mathcal{K}$, respectively, such that $\operatorname{Tr}\rho A=\alpha<+\infty$ and $\operatorname{Tr}\sigma B=\beta<+\infty$. It is easy to see that $C=A\otimes I_{\mathcal{K}}+I_{\mathcal{H}}\otimes B$ is an $\mathfrak{H}$-operator in the space $\mathcal{H}\otimes\mathcal{K}$ and

$$\operatorname{Tr}\omega C=\operatorname{Tr}\rho A+\operatorname{Tr}\sigma B=\alpha+\beta<+\infty \qquad \forall\,\omega\in\mathcal{C}(\rho,\sigma).$$

The remark before Lemma 3 implies the compactness of the set $\mathcal{C}(\rho,\sigma)$.

Proposition 2. *The set $\mathcal{P}_{\mathcal{A}}$ is a compact subset of $\mathcal{P}$ if and only if the set $\mathcal{A}$ is a compact subset of $\mathfrak{S}(\mathcal{H})$.*

Proof. Let the set $\mathcal{P}_{\mathcal{A}}$ be compact. The set $\mathcal{A}$ is the image of the set $\mathcal{P}_{\mathcal{A}}$ under the continuous mapping $\pi\mapsto\bar{\rho}(\pi)$, and hence it is compact.

Let the set $\mathcal{A}$ be compact. By Lemma 3 there exists an $\mathfrak{H}$-operator H such that $\operatorname{Tr}\rho H\leqq h$ for all ρ in $\mathcal{A}$. For arbitrary $\pi\in\mathcal{P}_{\mathcal{A}}$ we have

$$(10)\qquad \int_{\mathfrak{S}(\mathcal{H})}(\operatorname{Tr}\rho H)\,\pi(d\rho)=\operatorname{Tr}\left(\int_{\mathfrak{S}(\mathcal{H})}\rho\,\pi(d\rho)\,H\right)=\operatorname{Tr}\bar{\rho}(\pi)\,H\leqq h.$$

The existence of the integral on the left side and the first equality follows from the monotonous convergence theorem, since the function $\operatorname{Tr}\rho H$ is the limit of the nondecreasing sequence of continuous bounded functions $\operatorname{Tr}\rho Q_n H$ by (8).

Let $\mathcal{K}_\varepsilon=\{\rho\colon\ \operatorname{Tr}\rho H\leqq h\varepsilon^{-1}\}$. The set $\mathcal{K}_\varepsilon$ is compact for any ε. By (10) for any measure π in $\mathcal{P}_{\mathcal{A}}$ we have

$$(11)\quad \pi\big(\mathfrak{S}(\mathcal{H})\setminus\mathcal{K}_\varepsilon\big)=\int_{\mathfrak{S}(\mathcal{H})\setminus\mathcal{K}_\varepsilon}\pi(d\rho)\leqq\varepsilon h^{-1}\int_{\mathfrak{S}(\mathcal{H})\setminus\mathcal{K}_\varepsilon}(\operatorname{Tr}\rho H)\,\pi(d\rho)\leqq\varepsilon.$$

Compactness of the set $\mathcal{P}_{\mathcal{A}}$ follows from Prokhorov's theorem [12]. Proposition 2 is proved.

We will use the following notions, introduced in [17]. The sequence of ensembles $\{\pi_i^k,\rho_i^k\}$ with the averages $\bar{\rho}^k\in\mathcal{A}$ is called an *approximating sequence* if

$$\lim_{k\to+\infty}\chi_\Phi\big(\{\pi_i^k,\rho_i^k\}\big)=\overline{C}(\Phi;\,\mathcal{A}).$$

The state $\bar{\rho}\in\mathcal{A}$ is called an *optimal average state* for the $\mathcal{A}$-constrained channel Φ if it is a partial limit of a sequence of average states for some approximating sequence of ensembles. Compactness of the set $\mathcal{A}$ implies that the set of optimal average states is not empty.

Theorem. *Let $\mathcal{A}$ be a compact subset. If the restriction of the output entropy $H(\Phi(\rho))$ to the set $\mathcal{A}$ is continuous at an optimal average state $\bar{\rho}_0\in\mathcal{A}$, then there exists an optimal generalized ensemble π^* in $\mathcal{P}_{\mathcal{A}}$ such that* $\operatorname{supp}\pi^*\subseteq\operatorname{Extr}\mathfrak{S}(\mathcal{H})$ *and*

$$\overline{C}(\Phi;\,\mathcal{A})=\chi_\Phi(\pi^*)=\int_{\mathfrak{S}(\mathcal{H})}H\big(\Phi(\rho)\,\|\,\Phi\big(\bar{\rho}(\pi^*)\big)\big)\,\pi^*(d\rho).$$

Proof. We will show first that the functional

$$\pi\longmapsto\int_{\mathfrak{S}(\mathcal{H})}H\big(\Phi(\rho)\big)\,\pi(d\rho)$$

is well defined and lower semicontinuous on the set $\mathcal{P}$.

By Lemma 2 the function $H(\Phi(\rho))$ is a pointwise limit of the monotonously increasing sequences of functions

$$f_n(\rho) = \mathrm{Tr}\,\Big(\big(P_n\Phi(\rho)\,P_n\big)\big(I\log\mathrm{Tr}\,\big(P_n\Phi(\rho)\,P_n\big) - \log\big(P_n\Phi(\rho)\,P_n\big)\big)\Big),$$

which are continuous and bounded on $\mathfrak{S}(\mathcal{H})$. Hence the function $H(\Phi(\rho))$ is measurable and the monotonous convergence theorem implies

$$\int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho)\big)\,\pi(d\rho) = \lim_{n\to\infty}\int_{\mathfrak{S}(\mathcal{H})} f_n(\rho)\,\pi(d\rho).$$

The sequence of continuous functionals

$$\pi \longmapsto \int_{\mathfrak{S}(\mathcal{H})} f_n(\rho)\,\pi(d\rho)$$

is nondecreasing. Hence its pointwise limit is lower semicontinuous.

By the assumption the restriction of the function $H(\Phi(\rho))$ to the set $\mathcal{A}$ is continuous at some optimal average state $\bar\rho_0$. The continuity of the mapping $\pi \mapsto \bar\rho(\pi)$ implies that the restriction of the functional $\pi \mapsto H(\Phi(\bar\rho(\pi)))$ to the set $\mathcal{P}_{\mathcal{A}}$ is continuous at any point π_0 such that $\bar\rho(\pi_0) = \bar\rho_0$. Hence $H(\Phi(\bar\rho(\pi))) < +\infty$ for any point π in the intersection of $\mathcal{P}_{\mathcal{A}}$ with some neighborhood of π_0. For every such point π relation (6) holds. Therefore the restriction of the functional $\chi_\Phi(\pi)$ to the set $\mathcal{P}_{\mathcal{A}}$ is upper semicontinuous, and by Proposition 1 it is continuous at any point π_0 in $\mathcal{P}_{\mathcal{A}}$ such that $\bar\rho(\pi_0) = \bar\rho_0$.

Let $\{\pi_i^n, \rho_i^n\}$ be an approximating sequence of ensembles with the corresponding sequence of average states $\bar\rho^n$ converging to the state $\bar\rho_0$. Decomposing each state of the ensemble $\{\pi_i^n, \rho_i^n\}$ into a countable convex combination of pure states we obtain the sequence $\{\widehat\pi_j^n, \widehat\rho_j^n\}$ of generalized ensembles consisting of a countable number of pure states with the same sequence of the average states $\bar\rho^n$. Let $\widehat\pi^n$ be the sequence of measures ascribing value $\widehat\pi_j^n$ to the set $\{\widehat\rho_j^n\}$ for each j. It follows that

$$(12)\quad \chi_\Phi(\widehat\pi_n) = \sum_j \widehat\pi_j^n H\big(\Phi(\widehat\rho_j^n)\,\|\,\Phi(\bar\rho^n)\big) \geqq \sum_i \pi_i^n H\big(\Phi(\rho_i^n)\,\|\,\Phi(\bar\rho^n)\big) = \chi_\Phi\big(\{\pi_i^n, \rho_i^n\}\big),$$

where the inequality follows from the convexity of the relative entropy. By construction $\mathrm{supp}\,\widehat\pi^n \subseteq \mathrm{Extr}\,\mathfrak{S}(\mathcal{H})$ for each n. By Proposition 2 there exists a subsequence $\widehat\pi^{n_k}$, converging to some measure π^* in $\mathcal{P}_{\mathcal{A}}$. Since the set $\mathrm{Extr}\,\mathfrak{S}(\mathcal{H})$ of all pure states is a closed subset of $\mathfrak{S}(\mathcal{H})$[2], we have $\mathrm{supp}\,\pi^* \subseteq \mathrm{Extr}\,\mathfrak{S}(\mathcal{H})$ due to Theorem 6.1 in [13]. It is clear that $\bar\rho(\pi^*) = \bar\rho_0$, and, hence, as shown above, the restriction of the functional $\chi_\Phi(\pi)$ to the set $\mathcal{P}_{\mathcal{A}}$ is continuous at the point π^*. Thus the approximating property of the sequence $\{\pi_i^n, \rho_i^n\}$ and (12) imply

$$\overline{C}(\Phi;\,\mathcal{A}) = \lim_{k\to\infty}\chi_\Phi\big(\{\pi_i^{n_k},\,\rho_i^{n_k}\}\big) \leqq \lim_{k\to\infty}\chi_\Phi(\widehat\pi_{n_k}) = \chi_\Phi(\pi^*).$$

Since the converse inequality follows from Corollary 1, we obtain $\overline{C}(\Phi;\mathcal{A}) = \chi_\Phi(\pi^*)$, which means that the measure π^* is an optimal generalized ensemble for the $\mathcal{A}$-constrained channel Φ. The theorem is proved.

[2]The set $\mathrm{Extr}\,\mathfrak{S}(\mathcal{H})$ is described by the inequality $H(\rho) \leqq 0$, and due to the lower semicontinuity of the quantum entropy it is closed.

COROLLARY 2. *For the arbitrary state ρ_0 with $H(\Phi(\rho_0)) < +\infty$ there exists a generalized ensemble*[3] π_0 *such that* $\bar{\rho}(\pi_0) = \rho_0$ *and*

$$\chi_\Phi(\rho_0) \equiv \sup_{\sum_i \pi_i \rho_i = \rho_0} \chi_\Phi\bigl(\{\pi_i, \rho_i\}\bigr) = \int_{\mathfrak{S}(\mathcal{H})} H\bigl(\Phi(\rho) \,\|\, \Phi(\rho_0)\bigr)\, \pi_0(d\rho).$$

Proof. It is sufficient to note that the conditions of the theorem hold trivially for $\mathcal{A} = \{\rho_0\}$.

In the finite-dimensional case we obviously have

$$\overline{C}(\Phi; \mathcal{A}) = \chi_\Phi(\bar{\rho}), \tag{13}$$

where $\bar{\rho}$ is the average state of any optimal ensemble. The generalization of this relation to the infinite-dimensional case is closely connected to the question of the existence of the optimal generalized ensemble.

COROLLARY 3. *If an optimal generalized ensemble for the $\mathcal{A}$-constrained channel Φ exists, then equality* (13) *holds for some optimal average state $\bar{\rho}$ for the $\mathcal{A}$-constrained channel Φ.*

If equality (13) *holds for some optimal average state $\bar{\rho}$ for the $\mathcal{A}$-constrained channel Φ with $H(\Phi(\bar{\rho})) < +\infty$, then there exists an optimal generalized ensemble for the $\mathcal{A}$-constrained channel Φ.*

Proof. The first assertion is obvious while the second one follows from Corollary 2.

Remark. The continuity condition in the theorem is essential, as is shown in Appendix C. It is possible to show that this condition holds automatically if the set $\mathcal{A}$ is convex with a finite number of extreme points with finite output entropy. We conjecture that this condition holds for the arbitrary convex compact set $\mathcal{A}$ due to the special properties of optimal average states in this case, considered in [17].

PROPOSITION 3. *Let $\mathcal{A}$ be a compact set and H' be an $\mathfrak{H}$-operator in the space $\mathcal{H}'$ such that*

$$\operatorname{Tr}\exp(-\beta H') < +\infty \qquad \text{for all} \quad \beta > 0 \tag{14}$$

and $\operatorname{Tr}\Phi(\rho)\,H' \leqq h'$ *for all $\rho \in \mathcal{A}$. Then there exists an optimal generalized ensemble for the $\mathcal{A}$-constrained channel Φ.*

Proof. We show that under the conditions of the lemma the restriction of the output entropy $H(\Phi(\rho))$ to the set $\mathcal{A}$ is continuous, which implies validity of the conditions of the theorem.

Let $\rho'_\beta = (\operatorname{Tr}\exp(-\beta H'))^{-1}\exp(-\beta H')$ be a state in $\mathfrak{S}(\mathcal{H}')$. For arbitrary ρ in $\mathcal{A}$ we have

$$H\bigl(\Phi(\rho) \,\|\, \rho'_\beta\bigr) = -H\bigl(\Phi(\rho)\bigr) + \beta \operatorname{Tr}\Phi(\rho)\,H' + \log \operatorname{Tr}\exp(-\beta H'). \tag{15}$$

Let $\{\rho_n\}$ be an arbitrary sequence of states in $\mathcal{A}$ converging to the state ρ. By using (15) and the lower semicontinuity of the relative entropy we obtain

$$\begin{aligned}\limsup_{n\to\infty} H\bigl(\Phi(\rho_n)\bigr) &= H\bigl(\Phi(\rho)\bigr) + H\bigl(\Phi(\rho) \,\|\, \rho'_\beta\bigr) - \liminf_{n\to\infty} H\bigl(\Phi(\rho_n) \,\|\, \rho'_\beta\bigr) \\ &\quad + \limsup_{n\to\infty} \beta \operatorname{Tr}\Phi(\rho_n)\,H' - \beta \operatorname{Tr}\Phi(\rho)\,H' \leqq H\bigl(\Phi(\rho)\bigr) + \beta h'.\end{aligned}$$

[3] In what follows we can consider the generalized ensembles as measures supported by the set of pure states.

By letting β in the above inequality tend to zero we can establish the upper semicontinuity of the restriction of the function $H(\Phi(\rho))$ to the set $\mathcal{A}$. The lower semicontinuity of this function follows from the lower semicontinuity of the entropy [19]. Hence the restriction of the function $H(\Phi(\rho))$ to the set $\mathcal{A}$ is continuous. Proposition 3 is proved.

The condition of Proposition 3 is fulfilled for Gaussian channels with the power constraint of the form (9) where $H = R^T \epsilon R$ is the many-mode oscillator Hamiltonian with nondegenerate energy matrix ϵ, and R are the canonical variables of the system. We give a brief sketch of the argument which can be made rigorous by taking care of the unboundedness of the canonical variables. Indeed, let

$$R' = KR + K_E R_E$$

be the equation of the channel in the Heisenberg picture, where R_E are the canonical variables of the environment which is in the Gaussian state with zero mean and the correlation matrix α_E [9]. Taking $H' = c[R^T R - I\alpha_E K_E^T K_E]$, where denotes the trace of a matrix, we have $\Phi^*(H') = cR^T K^T KR$, and we can always choose a positive c such that $\Phi^*(H') \leqq H$. Moreover, H' satisfies condition (14). Thus the conditions of Proposition 3 can be fulfilled in this case.

CONJECTURE. *For an arbitrary Gaussian channel with the power constraint an optimal generalized ensemble is given by a Gaussian measure supported by the set of pure Gaussian states with arbitrary mean and a fixed correlation matrix.*

This conjecture was stated in [9] for the attenuation/amplification channel with classical noise. For the case of a pure attenuation channel characterized by the property of zero minimal output entropy the validity of this conjecture was established in [4]. Note also that the classical analogue of the above conjecture is the assertion that the optimal input distribution for a Gaussian channel with quadratic input constraint is Gaussian.

5. Convex constraints. In the case of a convex constraint set there are further special properties, such as the uniqueness of the output of an optimal average state; see [17]. The following lemma is a generalization of Donald's identity [3].

LEMMA 4. *For an arbitrary measure π in $\mathcal{P}$ and an arbitrary state σ in $\mathfrak{S}(\mathcal{H})$ the following identity holds*:

$$\int_{\mathfrak{S}(\mathcal{H})} H(\rho\,\|\,\sigma)\,\pi(d\rho) = \int_{\mathfrak{S}(\mathcal{H})} H\big(\rho\,\|\,\bar{\rho}(\pi)\big)\,\pi(d\rho) + H\big(\bar{\rho}(\pi)\,\|\,\sigma\big). \tag{16}$$

Proof. We first notice that in the finite-dimensional case Donald's identity

$$\sum_i \pi_i H(\rho_i\,\|\,\sigma) = \sum_i \pi_i H\big(\rho_i\,\|\,\bar{\rho}(\pi)\big) + H\big(\bar{\rho}(\pi)\,\|\,\sigma\big)$$

holds for not necessarily normalized positive operators with the generalized definition of the relative entropy (2). This can obviously be extended to generalized ensembles in finite-dimensional Hilbert space, giving (16) for this case. Thus this relation holds for the operators $P_n\rho P_n, P_n\sigma P_n$, where P_n is an arbitrary sequence of finite projectors increasing to $I_{\mathcal{H}}$. Passing to the limit as $n \to \infty$ and referring to the monotonous convergence theorem, we obtain (16) in the infinite-dimensional case. Lemma 4 is proved.

The following proposition is a generalization of the "maximal distance property" [10, Proposition 1].

PROPOSITION 4. *Let $\mathcal{A}$ be a convex subset of $\mathfrak{S}(\mathcal{H})$. A measure $\pi \in \mathcal{P}_{\mathcal{A}}$ is an optimal generalized ensemble for the $\mathcal{A}$-constrained channel Φ if and only if*

$$\int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi))\big)\, \mu(d\rho) \leqq \int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi))\big)\, \pi(d\rho) = \chi_{\Phi}(\pi) \tag{17}$$

for the arbitrary measure $\mu \in \mathcal{P}_{\mathcal{A}}$.

Proof. Let inequality (17) hold for the arbitrary measure $\mu \in \mathcal{P}_{\mathcal{A}}$. By Lemma 4 we have

$$\begin{aligned}\chi_{\Phi}(\mu) &\leqq \int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\mu))\big)\, \mu(d\rho) + H\Big(\Phi(\bar{\rho}(\mu)) \,\|\, \Phi(\bar{\rho}(\pi))\Big) \\ &= \int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi))\big)\, \mu(d\rho) \leqq \chi_{\Phi}(\pi),\end{aligned}$$

which implies the optimality of the measure π.

Conversely, let π be an optimal generalized ensemble for the $\mathcal{A}$-constrained channel Φ and let μ be an arbitrary measure in $\mathcal{P}_{\mathcal{A}}$. By the convexity of the set $\mathcal{A}$ the measure $\pi_{\eta} = \eta\mu + (1-\eta)\,\pi$ is also in $\mathcal{P}_{\mathcal{A}}$ for arbitrary $\eta \in (0,1)$. By using Lemma 4 we obtain

$$\begin{aligned}\chi_{\Phi}(\pi) \geqq \chi_{\Phi}(\pi_{\eta}) &= \int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi_{\eta}))\big)\, \pi_{\eta}(d\rho) \\ &= \eta \int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi_{\eta}))\big)\, \mu(d\rho) \\ &\quad + (1-\eta)\,\chi_{\Phi}(\pi) + (1-\eta)\, H\big(\bar{\rho}(\pi) \,\|\, \bar{\rho}(\pi_{\eta})\big).\end{aligned}$$

By the nonnegativity of the relative entropy we have

$$\int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi_{\eta}))\big)\, \mu(d\rho) \leqq \chi_{\Phi}(\pi). \tag{18}$$

By using the lower semicontinuity of the relative entropy we obtain

$$\begin{aligned}\liminf_{\eta\to 0} \int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi_{\eta}))\big)\, \mu(d\rho) &\geqq \int_{\mathfrak{S}(\mathcal{H})} \liminf_{\eta\to 0} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi_{\eta}))\big)\, \mu(d\rho) \\ &\geqq \int_{\mathfrak{S}(\mathcal{H})} H\big(\Phi(\rho) \,\|\, \Phi(\bar{\rho}(\pi))\big)\, \mu(d\rho).\end{aligned}$$

Proposition 4 is proved.

Appendix A.

THEOREM (see [1]). *Let ρ_n be a sequence of positive trace class operators converging to ρ in the weak operator topology and such that $\lim_{n\to\infty} \operatorname{Tr}\rho_n = \operatorname{Tr}\rho$. Then the sequence ρ_n converges to ρ in the trace norm.*

This implies that a sequence of quantum states converging to a state in the weak operator topology converges to it in the trace norm. One can consider this as a noncommutative generalization of the fact that weak convergence of probability distributions on a discrete probability space implies l_1 convergence. By using this theorem we can amplify to the trace norm topology the compactness criterion given in [14] under the name "the noncommutative Prokhorov theorem."

THE COMPACTNESS CRITERION. *A trace norm closed subset* $\mathcal{K}$ *of* $\mathfrak{S}(\mathcal{H})$ *is compact in the trace norm topology if and only if for arbitrary* $\varepsilon > 0$ *there exists a finite rank projector* P_ε *such that* $\operatorname{Tr} P_\varepsilon \rho > 1 - \varepsilon$ *for all* $\rho \in \mathcal{K}$.

Proof. Let $\mathcal{K}$ be a compact subset of $\mathfrak{S}(\mathcal{H})$. Suppose that there is $\varepsilon > 0$ such that for an arbitrary finite rank projector P there exists a state $\rho \in \mathcal{K}$ such that $\operatorname{Tr} P\rho \leqq 1 - \varepsilon$. Let P_n be a sequence of finite rank projectors in $\mathcal{H}$ monotonously converging to the identity operator $I_\mathcal{H}$ in the weak operator topology and ρ_n be the corresponding sequence of states in $\mathcal{K}$. By the compactness of $\mathcal{K}$ there exists a subsequence ρ_{n_k} converging to a state $\rho_* \in \mathcal{K}$. By construction

$$\operatorname{Tr} P_{n_l}\rho_{n_k} \leqq \operatorname{Tr} P_{n_k}\rho_{n_k} \leqq 1 - \varepsilon \qquad \text{for all} \quad k > l.$$

Hence

$$\operatorname{Tr} \rho_* = \lim_{l\to+\infty} \operatorname{Tr} P_{n_l}\rho_* = \lim_{l\to+\infty} \lim_{k\to+\infty} \operatorname{Tr} P_{n_l}\rho_{n_k} \leqq 1 - \varepsilon,$$

which contradicts the fact that $\rho_* \in \mathcal{K} \subseteq \mathfrak{S}(\mathcal{H})$.

Conversely, let $\mathcal{K}$ be a subset of $\mathfrak{S}(\mathcal{H})$ satisfying the criterion. Let ρ_n be an arbitrary sequence in $\mathcal{K}$. Since the unit ball in $\mathfrak{B}(\mathcal{H})$ is compact in the weak operator topology, there exists a subsequence ρ_{n_k} converging to a positive operator ρ_* in this topology. We have

$$\operatorname{Tr} \rho_* \leqq \liminf_{k\to\infty} \operatorname{Tr} \rho_{n_k} = 1,$$

therefore to prove that ρ_* is a state it is sufficient to show that $\operatorname{Tr} \rho_* \geqq 1$. Let $\varepsilon > 0$ and P_ε be the corresponding projector. Then

$$\operatorname{Tr} \rho_* \geqq \operatorname{Tr} P_\varepsilon \rho_* = \lim_{k\to\infty} \operatorname{Tr} P_\varepsilon \rho_{n_k} > 1 - \varepsilon,$$

where the equality follows from the finite dimensionality of the space $P_\varepsilon(\mathcal{H})$. Thus ρ_* is a state. The theorem given above implies that the subsequence ρ_{n_k} converges to the state ρ_* in the trace norm. Thus the set $\mathcal{K}$ is trace norm compact.

Appendix B.

Proof of Lemma 1. We first notice that $\operatorname{supp} \pi \subseteq U$, where U is a closed convex subset of $\mathfrak{S}(\mathcal{H})$, implies

$$\bar{\rho}(\pi) \in U. \tag{19}$$

This is obvious for an arbitrary measure π with finite support. By Theorem 6.3 in [13] the set of such measures is dense in $\mathcal{P}$. The continuity of the mapping $\pi \mapsto \bar{\rho}(\pi)$ completes the proof of (19).

Now let π be an arbitrary measure in $\mathcal{P}$. Since $\mathfrak{S}(\mathcal{H})$ is separable, for each $n \in \mathbf{N}$ there exists a sequence $\{A_i^n\}$ of Borel sets of diameters less than $1/n$ such that $\mathfrak{S}(\mathcal{H}) = \bigcup_i A_i^n$, $A_i^n \cap A_j^n = \varnothing$ provided $j \neq i$. Let $m = m(n)$ be a number such that $\sum_{i=m+1}^{+\infty} \pi(A_i^n) < 1/n$. Consider the finite collection of Borel sets $\{\widehat{A}_i^n\}_{i=1}^{m+1}$, where $\widehat{A}_i^n = A_i^n$ for all $i = 1, \dots, m$ and $\widehat{A}_{m+1}^n = \bigcup_{i=m+1}^{+\infty} A_i^n$. We have

$$\bar{\rho}(\pi) = \sum_{i=1}^{m+1} \int_{\widehat{A}_i^n} \rho\, \pi(d\rho) = \sum_{i=1}^{m+1} \pi_i^n \rho_i^n, \tag{20}$$

where $\pi_i^n = \mathrm{Tr} \int_{\widehat{A}_i^n} \rho\, \pi(d\rho) = \pi(\widehat{A}_i^n)$ and $\rho_i^n = (\pi(\widehat{A}_i^n))^{-1} \int_{\widehat{A}_i^n} \rho\, \pi(d\rho)$ (without loss of generality we assume $\pi_i^n > 0$). Let π^n be the probability measure on $\mathfrak{S}(\mathcal{H})$ ascribing the value π_i^n to the set $\{\rho_i^n\}$. Equality (20) implies $\bar{\rho}(\pi^n) = \bar{\rho}(\pi)$. Since the measure π^n has finite support for each n, to prove the assertion of the lemma it is sufficient to show weak convergence of the sequence of measures π^n to the measure π. By Theorem 6.1 in [13], to establish the above convergence it is sufficient to show that

$$\lim_{n\to+\infty} \int_{\mathfrak{S}(\mathcal{H})} f(\rho)\, \pi^n(d\rho) = \int_{\mathfrak{S}(\mathcal{H})} f(\rho)\, \pi(d\rho)$$

for the arbitrary bounded uniformly continuous function $f(\rho)$ on $\mathfrak{S}(\mathcal{H})$. Let $M_f = \sup_{\rho\in\mathfrak{S}(\mathcal{H})} |f(\rho)|$. For arbitrary $\varepsilon > 0$ let n_ε be such that $\varepsilon n_\varepsilon > 2M_f$ and

$$\sup_{\rho\in U(n_\varepsilon)} f(\rho) - \inf_{\rho\in U(n_\varepsilon)} f(\rho) < \varepsilon$$

for the arbitrary closed ball $U(n_\varepsilon)$ of diameter $1/n_\varepsilon$. Let $n \geqq n_\varepsilon$. By construction the set $\widehat{A}_i^n$ is contained in some ball $U_i(n)$ for each $i = 1,\dots,m$. By (19) the state ρ_i^n lies in the same ball $U_i(n)$. Hence we have

$$\left| \int_{\mathfrak{S}(\mathcal{H})} f(\rho)\, \pi^n(d\rho) - \int_{\mathfrak{S}(\mathcal{H})} f(\rho)\, \pi(d\rho) \right| \leqq \sum_{i=1}^{m+1} \int_{\widehat{A}_i^n} \left| f(\rho) - f(\rho_i) \right| \pi(d\rho)$$
$$\leqq \varepsilon \sum_{i=1}^{m} \pi(\widehat{A}_i^n) + 2M_f \pi(\widehat{A}_{m+1}^n) < 2\varepsilon \qquad \text{for all } n \geqq n_\varepsilon.$$

Appendix C. Example of a channel without optimal generalized ensembles. We give an example of a classical channel, which can be extended to the quantum one in a standard way. Consider the abelian von Neumann algebra l_∞ and its predual l_1, which can be considered as spaces of diagonal operators in the separable Hilbert space l_2. Let Φ be the identity channel in l_1. Consider the sequence of states (probability distributions)

$$\rho_n = \{1 - q_n, \underbrace{n^{-1}q_n,\, n^{-1}q_n, \dots, n^{-1}q_n}_{n}, 0, 0, \dots\},$$

where q_n is a sequence in $[0,1]$, which will be defined below. Note that in this case $\chi_\Phi(\rho_n) = H(\rho_n) = h_2(q_n) + q_n \log n$, where $h_2(x) = -x\log x - (1-x)\log(1-x)$. We will show later that there exists a sequence q_n such that $\lim_{n\to+\infty} q_n = 0$, while the corresponding sequence $\chi_\Phi(\rho_n) = H(\rho_n)$ monotonously increases to 1. Let q_n be such a sequence and let $\mathcal{A}$ be the closure of the sequence ρ_n, which obviously consists of states ρ_n and the pure state $\rho_* = \lim_{n\to+\infty} \rho_n = \{1, 0, 0, \dots\}$. By the definition and the monotonicity $\overline{C}(\Phi; \mathcal{A}) = \lim_{n\to+\infty} \chi_\Phi(\rho_n) = 1$, while ρ_* is the only optimal average state for the $\mathcal{A}$-constrained channel Φ and $\chi_\Phi(\rho_*) = H(\rho_*) = 0$. Thus we have $\overline{C}(\Phi; \mathcal{A}) > \chi_\Phi(\rho_*)$ and Corollary 3 implies that there is no optimal ensemble for the $\mathcal{A}$-constrained channel Φ.

Let us construct the sequence q_n with the above properties. Consider the strictly increasing function $f(x) = x(1 - \log x)$ on $[0,1]$. It easy to see that $f'(x) = -\log x$ and $f([0,1]) = [0,1]$. Let f^{-1} be the converse function and $g(x) = xf^{-1}((\log 2)/x)$ for all $x \geqq 1$. Note that the function $g(x)$ is implicitly defined by the equation

$$g\left(1 - \log\left(\frac{g}{x}\right)\right) = \log 2. \tag{21}$$

Using this fact, it is easy to see that the function $g(x)$ satisfies the following differential equation

$$\log\left(\frac{g}{x}\right) g' = \frac{g}{x}. \tag{22}$$

Since $g(x)/x = f^{-1}((\log 2)/x)$, we have $g(x)/x \in [0,1]$. This fact, with (21) and (22), implies $g(x) \in [0,1]$, $\lim_{x\to+\infty} g(x) = 0$, and $g'(x) < 0$. Consider the function $H(x) = h_2(g(x)) + g(x)\log x$. By the above-mentioned properties of the function $g(x)$, (21) and (22) we obtain $\lim_{x\to+\infty} H(x) = (\log 2)^{-1} \lim_{x\to+\infty} g(x)\log x = 1$ and

$$\begin{aligned} H'(x) &= (\log 2)^{-1}\left(g'(x)\,\log\big(1-g(x)\big) - g'(x)\,\log g(x) + g'(x)\,\log x + \frac{g(x)}{x}\right) \\ &= (\log 2)^{-1} g'(x)\,\log(1-g(x)) > 0 \qquad \forall\, x > 1. \end{aligned}$$

It follows that $H(x)$ is an increasing function on $[1,+\infty)$, tending to 1 at infinity. Setting $q_n = g(n)$ we obtain the sequence with the desired properties.

REFERENCES

[1] G. F. Dell'Antonio, *On the limits of sequences of normal states*, Comm. Pure Appl. Math., 20 (1967), pp. 413–429.

[2] P. Billingsley, *Convergence of Probability Measures*, John Wiley, New York, 1968.

[3] M. J. Donald, *Further results on the relative entropy*, Math. Proc. Cambridge Philos. Soc., 101 (1987), pp. 363–374.

[4] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, *Classical capacity of the lossy bosonic channels: The exact solution*, Phys. Rev. Lett., 92 (2004), 027902; LANL Report quant-ph/0308012.

[5] E. Hille and R. S. Phillips, *Functional Analysis and Semi-groups*, AMS, Providence, RI, 1957.

[6] A. S. Holevo, *An Introduction to Quantum Information Theory*, MCCME, Moscow, 2002 (in Russian).

[7] A. S. Holevo, *Statistical Structure of Quantum Theory*, Springer-Verlag, New York, 2001.

[8] A. S. Holevo, *Entanglement-assisted capacities of constrained quantum channels*, Theory Probab. Appl., 48 (2003), pp. 243–255.

[9] A. S. Holevo and R. F. Werner, *Evaluating capacities of bosonic Gaussian channels*, Phys. Rev. A, 63 (2001), 032312.

[10] A. S. Holevo and M. E. Shirokov, *On Shor's channel extension and constrained channels*, Comm. Math. Phys., 249 (2004), pp. 417–430.

[11] G. Lindblad, *Expectations and entropy inequalities for finite quantum systems*, Comm. Math. Phys., 39 (1974), pp. 111–119.

[12] Yu. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Theory Probab. Appl., 1 (1956), pp. 157–214.

[13] K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York, London, 1967.

[14] T. A. Sarymsakov, *Introduction to Quantum Probability Theory*, FAN, Tashkent, 1985 (in Russian).

[15] B. Schumacher and M. Westmoreland, *Optimal Signal Ensembles*, Tech. Report quant-ph/9912122, Los Alamos National Lab., Los Alamos, NM, 1999.

[16] A. Serafini, J. Eisert, and M. M. Wolf, *Multiplicativity of Maximal Output Purities of Gaussian Channels under Gaussian Inputs*, Tech. Report quant-ph/0406065, Los Alamos National Lab., Los Alamos, NM, 2004.

[17] M. E. Shirokov, *The Holevo Capacity of Infinite-Dimensional Channels*, LANL Report quant-ph/0408009, Los Alamos National Lab., Los Alamos, NM, 2004.

[18] P. W. Shor, *Equivalence of additivity questions in quantum information theory*, Comm. Math. Phys., 246 (2004), pp. 453–472.

[19] A. Wehrl, *General properties of entropy*, Rev. Modern Phys., 50 (1978), pp. 221–260.