Tight continuity bounds for the quantum conditional mutual information, for the Holevo quantity and for capacities of a channel

M.E. Shirokov*

Abstract

First we consider Fannes’ type and Winter’s type tight continuity bounds for the quantum conditional mutual information and their specifications for states of special types.

Then we analyse continuity of the Holevo quantity with respect to two nonequivalent metrics on the set of ensembles of quantum states. We show that the Holevo quantity is continuous on the set of all ensembles of \( m \) states with respect to both metrics if either \( m \) or the dimension of underlying Hilbert space is finite and obtain Fannes’ type tight continuity bounds for the Holevo quantity in this case.

In general case conditions for local continuity of the Holevo quantity and their corollaries (preserving local continuity under quantum channels, stability with respect to perturbation of states) are considered. Winter’s type tight continuity bound for the Holevo quantity under the energy constraint on the average state of ensembles is obtained and applied to the system of quantum oscillators.

The above results are used to obtain tight and close-to-tight continuity bounds for basic capacities of finite-dimensional channels (refining the Leung-Smith continuity bounds) and for classical capacities of infinite-dimensional channels with energy constraints.

*Steklov Mathematical Institute, RAS, Moscow, email:msh@mi.ras.ru
## Contents

1 Introduction 2

2 Preliminaries 4

3 Tight continuity bounds for the quantum conditional mutual information 6
   3.1 Fannes’ type continuity bounds for $I(A:B|C)$ 8
   3.2 Winter’s type continuity bound for $I(A:B|C)$ 11
   3.3 Continuity bound for the function $\Phi \mapsto I(B^n:D|C)_{\Phi^{\otimes n}(\rho)}$ 13

4 On continuity of the Holevo quantity 16
   4.1 Two nonequivalent metrics on the set of quantum ensembles 17
   4.2 The case of global continuity 19
   4.3 General case 22

5 Applications 27
   5.1 Tight continuity bounds for the Holevo capacity and for the entanglement-assisted classical capacity of a quantum channel 27
   5.2 Refinement of the Leung-Smith continuity bounds for classical and quantum capacities of a channel 29
   5.3 Continuity bounds for classical capacities of infinite-dimensional channels with energy constraints 31

## 1 Introduction

A quantitative analysis of continuity of basis characteristics of quantum systems and channels is a necessary technical tool in study of their information properties. It suffices to mention that the famous Fannes continuity bound for the von Neumann entropy and the Alicki-Fannes continuity bound for the conditional entropy are essentially used in the proofs of several important results of quantum information theory [13, 23, 31]. During the last decade many papers devoted to finding continuity bounds (estimates for variation) for different quantities have been appeared (see [2, 3, 4, 18, 26, 32] and the references therein).

Although in many applications a structure of a continuity bound of a given quantity is more important than concrete values of its coefficients, a
task of finding optimal values of these coefficients seems interesting from the both mathematical and physical points of view. This task can be formulated as a problem of finding so called "tight" continuity bound, i.e. \( \varepsilon \)-achievable estimates for variations of a given quantity. The most known decision of this problem is the sharpest continuity bound for the von Neumann entropy obtained by Audenaert [2] (it refines the Fannes continuity bound mentioned above). Other result in this direction is the tight bound for the relative entropy difference via the entropy difference obtained by Reeb and Wolf [26]. Recently Winter presented tight continuity bound for the conditional entropy (improving the Alicki-Fannes continuity bound) and asymptotically tight continuity bounds for the entropy and for the conditional entropy in infinite-dimensional systems under energy constraint [32]. By using Winter’s technique a tight continuity bound for quantum conditional mutual information in infinite-dimensional tripartite systems under the energy constraint on one subsystem is obtained in [29, the Appendix].

In this paper we specify Fannes’ type and Winter’s type tight continuity bounds for the quantum conditional mutual information (obtained respectively in [28] and [29]). Then, by using the Leung-Smith telescopic trick from [18] tight continuity bounds of the both types for the output quantum conditional mutual information for \( n \)-tensor power of a channel are obtained.

We analyse continuity properties of the Holevo quantity with respect to two nonequivalent metrics \( D_0 \) and \( D_* \) on the set of ensembles of quantum states. The metric \( D_0 \) is a trace norm distance between ensembles considered as ordered collections of states, the metric \( D_* \) is a factorization of \( D_0 \) obtained by identification of all ensembles corresponding to the same probability measure on the set of quantum states. It is shown that \( D_* \) coincides with the EHS-distance between ensembles introduced by Oreshkov and Cal-samiglia in [24] and that \( D_* \) generates the weak convergence topology on the set of all ensembles considered as probability measures.

We show that the Holevo quantity is continuous on the set of all ensembles of \( m \) states with respect to the metrics \( D_0 \) and \( D_* \) if either \( m \) or the dimension of underlying Hilbert space is finite and obtain Fannes’ type tight continuity bounds for the Holevo quantity with respect to both metrics in this case.

In general case conditions for local continuity of the Holevo quantity with respect to the metrics \( D_0 \) and \( D_* \) and their corollaries (preserving of local continuity under quantum channels, stability with respect to perturbation of states) are considered. Winter’s type tight continuity bound for the Holevo quantity under the energy constraint on the average state of ensembles is
obtained and applied to the system of quantum oscillators.

The above results are used to obtain tight and close-to-tight continuity bounds for basic capacities of channels with finite-dimensional output (essentially refining the Leung-Smith continuity bounds from [18]) and for classical capacities of infinite-dimensional channels with energy constraints.

2 Preliminaries

Let $\mathcal{H}$ be a finite-dimensional or separable infinite-dimensional Hilbert space, $\mathfrak{B}(\mathcal{H})$ the algebra of all bounded operators with the operator norm $\| \cdot \|$ and $\mathfrak{T}(\mathcal{H})$ the Banach space of all trace-class operators in $\mathcal{H}$ with the trace norm $\| \cdot \|_1$. Let $\mathcal{S}(\mathcal{H})$ be the set of quantum states (positive operators in $\mathfrak{T}(\mathcal{H})$ with unit trace) [13, 23, 31].

Denote by $I_{\mathcal{H}}$ the unit operator in a Hilbert space $\mathcal{H}$ and by $\text{Id}_{\mathcal{H}}$ the identity transformation of the Banach space $\mathfrak{T}(\mathcal{H})$.

A finite or countable collection $\{\rho_i\}$ of states with a probability distribution $\{p_i\}$ is conventionally called ensemble and denoted $\{p_i, \rho_i\}$. The state $\bar{\rho} = \sum_i p_i \rho_i$ is called average state of this ensemble.

If quantum systems $A$ and $B$ are described by Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ then the bipartite system $AB$ is described by the tensor product of these spaces, i.e. $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. A state in $\mathcal{S}(\mathcal{H}_{AB})$ is denoted $\omega_{AB}$, its marginal states $\text{Tr}_{\mathcal{H}_B} \omega_{AB}$ and $\text{Tr}_{\mathcal{H}_A} \omega_{AB}$ are denoted respectively $\omega_A$ and $\omega_B$. In this paper a special role is played by so called $qc$-states having the form

$$\omega_{AB} = \sum_{i=1}^{m} p_i \rho_i \otimes |i\rangle \langle i|, \quad (1)$$

where $\{\rho_i, \rho_i\}_{i=1}^{m}$ is an ensemble of $m \leq +\infty$ quantum states in $\mathcal{S}(\mathcal{H}_A)$ and $\{|i\rangle\}_{i=1}^{m}$ is an orthonormal basis in $\mathcal{H}_B$.

The von Neumann entropy $H(\rho) = \text{Tr} \eta(\rho)$ of a state $\rho \in \mathcal{S}(\mathcal{H})$, where $\eta(x) = -x \log x$, is a concave nonnegative lower semicontinuous function on $\mathcal{S}(\mathcal{H})$, it is continuous if and only if $\dim \mathcal{H} < +\infty$ [22, 30].

The concavity of the von Neumann entropy is supplemented by the inequality

$$H(\lambda \rho + (1-\lambda)\sigma) \leq \lambda H(\rho) + (1-\lambda)H(\sigma) + h_2(\lambda), \quad \lambda \in [0, 1], \quad (2)$$

where $h_2(\lambda) = \eta(\lambda) + \eta(1-\lambda)$, valid for any states $\rho$ and $\sigma$ [23].
Audenaert obtained in [2] the sharpest continuity bound for the von Neumann entropy:

$$|H(\rho) - H(\sigma)| \leq \varepsilon \log(d - 1) + h_2(\varepsilon)$$

(3)

for any $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ such that $\varepsilon = \frac{1}{2}\|\rho - \sigma\|_1 \leq 1 - 1/d$, where $d = \dim \mathcal{H}$. This continuity bound is a refinement of the well known Fannes continuity bound [9].

The quantum conditional entropy

$$H(A|B)_\omega = H(\omega_{AB}) - H(\omega_B)$$

(4)

of a bipartite state $\omega_{AB}$ with finite marginal entropies is essentially used in analysis of quantum systems [13, 31]. The function $\omega_{AB} \mapsto H(A|B)_\omega$ is continuous on $\mathcal{S}(\mathcal{H}_{AB})$ if and only if $\dim \mathcal{H}_A < +\infty$.

The conditional entropy is concave and satisfies the following inequality

$$H(A|B)_{\lambda\rho+(1-\lambda)\sigma} \leq \lambda H(A|B)_\rho + (1 - \lambda)H(A|B)_\sigma + h_2(\lambda)$$

(5)

for any $\lambda \in (0, 1)$ and any states $\rho_{AB}$ and $\sigma_{AB}$. Inequality (5) follows from concavity of the entropy and inequality (2).

Winter proved in [32] the following refinement of the Alicki-Fannes continuity bound for the conditional entropy (obtained in [1]):

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq 2\varepsilon \log d + (1 + \varepsilon)h_2\left(\frac{\varepsilon}{1 + \varepsilon}\right)$$

(6)

for any states $\rho, \sigma \in \mathcal{S}(\mathcal{H}_{AB})$ such that $\varepsilon = \frac{1}{2}\|\rho - \sigma\|_1$, where $d = \dim \mathcal{H}_A$. He showed that this continuity bound is tight and that the factor 2 in (6) can be removed if $\rho$ and $\sigma$ are qc-states, i.e. states having form (1).

Winter also obtained asymptotically tight continuity bounds for the entropy and for the conditional entropy for infinite-dimensional quantum states with bounded energy (see details in [32]).

The quantum relative entropy for two states $\rho$ and $\sigma$ in $\mathcal{S}(\mathcal{H})$ is defined as follows

$$H(\rho \| \sigma) = \sum \langle i | \rho \log \rho - \rho \log \sigma | i \rangle,$$

1If $\dim \mathcal{H}_A < +\infty$ and $\dim \mathcal{H}_B = +\infty$ then formula (41) is not well defined for some states, but there is an alternative expression for $H(A|B)_\omega$ (derived from the below formula (18) with trivial $C$) which gives concave continuous function on $\mathcal{S}(\mathcal{H}_{AB})$ in this case [17].
where \( \{|i\}\} \) is the orthonormal basis of eigenvectors of the state \( \rho \) and it is assumed that \( H(\rho \| \sigma) = +\infty \) if \( \text{supp}\rho \) is not contained in \( \text{supp}\sigma \) [22].

Several continuity bounds for the relative entropy are proved by Audenaert and Eisert [3, 4]. Tight bound for the relative entropy difference expressed via the entropy difference is obtained by Reeb and Wolf [26].

A quantum channel \( \Phi \) from a system \( A \) to a system \( B \) is a completely positive trace preserving linear map \( \mathcal{T}(\mathcal{H}_A) \rightarrow \mathcal{T}(\mathcal{H}_B) \), where \( \mathcal{H}_A \) and \( \mathcal{H}_B \) are Hilbert spaces associated with the systems \( A \) and \( B \) [13, 23, 31].

Denote by \( \mathcal{F}(A,B) \) the set of all quantum channels from from a system \( A \) to a system \( B \). We will use two metrics on the set \( \mathcal{F}(A,B) \) induced respectively by the operator norm \( \|\Phi\| = \sup_{\rho \in \mathcal{T}(\mathcal{H}_A), \|\rho\|_1 = 1} \|\Phi(\rho)\|_1 \)

and by the diamond norm

\[
\|\Phi\|_\diamond = \sup_{\rho \in \mathcal{T}(\mathcal{H}_{AB}), \|\rho\|_1 = 1} \|\Phi \otimes \text{Id}_R(\rho)\|_1,
\]

of a map \( \Phi : \mathcal{T}(\mathcal{H}_A) \rightarrow \mathcal{T}(\mathcal{H}_B) \). The latter coincides with the norm of complete boundedness of the dual map \( \Phi^* : \mathcal{B}(\mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_A) \) to \( \Phi \) [13, 31].

3 Tight continuity bounds for the quantum conditional mutual information

The quantum mutual information of a bipartite state \( \omega_{AB} \) is defined as follows

\[
I(A:B)_\omega = H(\omega_{AB} \| \omega_A \otimes \omega_B) = H(\omega_A) + H(\omega_B) - H(\omega_{AB}), \quad (7)
\]

where the second expression is valid if \( H(\omega_{AB}) \) is finite [21].

Basic properties of the relative entropy show that \( \omega \mapsto I(A:B)_{\omega} \) is a lower semicontinuous function on the set \( \mathcal{S}(\mathcal{H}_{AB}) \) taking values in \([0, +\infty]\). It is well known that

\[
I(A:B)_\omega \leq 2 \min \{H(\omega_A), H(\omega_B)\} \quad (8)
\]
for any state \( \omega_{AB} \) and that
\[
I(A:B)_\omega \leq \min \{ H(\omega_A), H(\omega_B) \}
\]
(9) for any separable state \( \omega_{AB} \) \[19, 31\].

The quantum conditional mutual information of a state \( \omega_{ABC} \) of a tripartite finite-dimensional system is defined by
\[
I(A:B|C)_\omega = H(\omega_{AC}) + H(\omega_{BC}) - H(\omega_{ABC}) - H(\omega_C).
\]
(10)

This quantity plays important role in quantum information theory \[?, 31\], its nonnegativity is a basic result well known as strong subadditivity of von Neumann entropy \[20\]. If system \( C \) is trivial then (10) coincides with (7).

In infinite dimensions formula (10) may contain the uncertainty “\( \infty - \infty \)”. Nevertheless the conditional mutual information can be defined for any state \( \omega_{ABC} \) by one of the equivalent expressions
\[
I(A:B|C)_\omega = \sup_{P_A} [I(A:BC)_{Q_A\omega Q_A} - I(A:C)_{Q_A\omega Q_A}] , Q_A = P_A \otimes I_{BC},
\]
(11)

\[
I(A:B|C)_\omega = \sup_{P_B} [I(B:AC)_{Q_B\omega Q_B} - I(B:C)_{Q_B\omega Q_B}] , Q_B = P_B \otimes I_{AC},
\]
(12)

where the suprema are over all finite rank projectors \( P_A \in \mathcal{B}(\mathcal{H}_A) \) and \( P_B \in \mathcal{B}(\mathcal{H}_B) \) correspondingly and it is assumed that \( I(X:Y)_{Q_X\omega Q_X} = \lambda I(X:Y)_{\lambda^{-1}Q_X\omega Q_X} \), where \( \lambda = \text{Tr}Q_X\omega_{ABC} \) \[28\].

It is shown in \[\text{28, Th.2}\] that expressions (11) and (12) define the same lower semicontinuous function on the set \( \mathcal{S}(\mathcal{H}_{ABC}) \) possessing all basic properties of conditional mutual information valid in finite dimensions. In particular, the following relation (chain rule)
\[
I(X:YZ|C)_\omega = I(X:Y|C)_\omega + I(X:Z|YC)_\omega
\]
(13)

holds for any state \( \omega \) in \( \mathcal{S}(\mathcal{H}_{XYZC}) \) (with possible values \(+\infty\) in both sides). To prove (13) is suffices to note that it holds if the systems \( X, Y, Z \) and \( C \) are finite-dimensional and to apply the approximating property from Corollary 9 in \[\text{28}\].

If one of the marginal entropies \( H(\omega_A) \) and \( H(\omega_B) \) is finite then the conditional mutual information is given respectively by the explicit formula
\[
I(A:B|C)_\omega = I(A:BC)_\omega - I(A:C)_\omega,
\]
(14)

\[\text{2The correctness of these formulae follows from upper bound} \ [\text{8}].\]
and
\[ I(A:B|C)_\omega = I(B:AC)_\omega - I(B:C)_\omega. \] (15)

By applying upper bound (8) to expressions (14) and (15) we see that
\[ I(A:B|C)_\omega \leq 2 \min \{ H(\omega_A), H(\omega_B), H(\omega_{AC}), H(\omega_{BC}) \} \] (16)
for any state \( \omega_{ABC} \).

The quantum conditional mutual information is not concave or convex but the inequality
\[ |\lambda I(A:B|C)_\rho + (1 - \lambda) I(A:B|C)_\sigma - I(A:B|C)_{\lambda \rho + (1 - \lambda) \sigma} | \leq h_2(\lambda) \] (17)
holds for \( \lambda \in (0, 1) \) and any states \( \rho_{ABC}, \sigma_{ABC} \) with finite \( I(A:B|C)_\rho, I(A:B|C)_\sigma \). If \( \rho_{ABC}, \sigma_{ABC} \) are states with finite marginal entropies then (17) can be easily proved by noting that
\[ I(A:B|C)_\omega = H(A|C)_\omega - H(A|BC)_\omega, \] (18)
and by using concavity of the conditional entropy and inequality (5). The validity of inequality (17) for any states \( \rho_{ABC}, \sigma_{ABC} \) with finite conditional mutual information can be proved by approximation (using the second part of Theorem 2 in [28]).

### 3.1 Fannes’ type continuity bounds for \( I(A:B|C) \).

Property (17) makes it possible to directly apply Winter’s modification of the Alicki-Fannes technic (cf. [1, 32]) to the conditional mutual information.

**Proposition 1.** Let \( \rho_{ABC} \) and \( \sigma_{ABC} \) be states such that
\[ D \doteq \max \{ I(A:B|C)_{\tau_-}, I(A:B|C)_{\tau_+} \} < +\infty, \quad \text{where} \quad \tau_\pm = \frac{[\sigma - \rho]_\pm}{\text{Tr}[\sigma - \rho]_\pm}. \]

Then
\[ |I(A:B|C)_\rho - I(A:B|C)_\sigma| \leq D\varepsilon + 2g(\varepsilon), \] (19)
where \( \varepsilon = \frac{1}{2}\|\rho - \sigma\|_1 \) and \( g(\varepsilon) \doteq (1 + \varepsilon)h_2\left(\frac{\varepsilon}{1+\varepsilon}\right) = (1 + \varepsilon) \log(1 + \varepsilon) - \varepsilon \log \varepsilon \). \footnote{\([\omega]_+ \) and \([\omega]_- \) are respectively positive and negative parts of an operator \( \omega \).} \footnote{\( \varepsilon \) is involved in the expression for entropy of Gaussian states [13, Ch.12].}
If the states $\rho_X$ and $\sigma_X$, where $X$ is one of the subsystems $A, B, AC, BC$, are supported by some $d$-dimensional subspace of $\mathcal{H}_X$ then (19) holds with $D = 2 \log d$.

**Proof.** Following [32] introduce the state $\omega^* = (1 + \varepsilon)^{-1}(\rho + [\sigma - \rho]_+)$. Then

$$\frac{1}{1 + \varepsilon} \rho + \frac{\varepsilon}{1 + \varepsilon} \tau_+ = \omega^* = \frac{1}{1 + \varepsilon} \sigma + \frac{\varepsilon}{1 + \varepsilon} \tau_-,$$

where $\tau_+ = \varepsilon^{-1} [\sigma - \rho]_+$ and $\tau_- = \varepsilon^{-1} [\sigma - \rho]_-$ are states in $\mathcal{S}(\mathcal{H}_{ABC})$. By applying (17) to the above convex decompositions of $\omega^*$ we obtain

$$(1 - p) [I(A: B|C)_\rho - I(A: B|C)_\sigma] \leq p [I(A: B|C)_{\tau_-} - I(A: B|C)_{\tau_+}] + 2h_2(p)$$

and

$$(1 - p) [I(A: B|C)_{\sigma} - I(A: B|C)_\rho] \leq p [I(A: B|C)_{\tau_+} - I(A: B|C)_{\tau_-}] + 2h_2(p),$$

where $p = \frac{\varepsilon}{1 + \varepsilon}$. These inequalities and nonnegativity of $I(A: B|C)$ imply (19).

The last assertion of the proposition follows from the first one and upper bound (16), since the states $[\tau_+]_X$ are supported by the minimal subspace of $\mathcal{H}_X$ containing the supports of $\rho_X$ and $\sigma_X$. □

Proposition 4 implies the following refinement of Corollary 8 in [28].

**Corollary 1.** If $d \equiv \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\} < +\infty$ then

$$|I(A: B|C)_\rho - I(A: B|C)_\sigma| \leq 2\varepsilon \log d + 2g(\varepsilon) \quad (20)$$

for any states $\rho, \sigma$ in $\mathcal{S}(\mathcal{H}_{ABC})$, where $\varepsilon = \frac{1}{2} \| \rho - \sigma \|_1$. Continuity bound (20) is tight even for trivial $C$, i.e. in the case $I(A: B|C) = I(A: B)$.

**Proof.** Continuity bound (20) directly follows from Proposition 4.

The tightness of this bound with trivial $C$ can be shown by using the example from [32, Remark 3]. Let $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$, $\rho_{AB}$ be a maximally entangled pure state and $\sigma_{AB} = (1 - \varepsilon)\rho_{AB} + \frac{\varepsilon}{d - 1}(I_{AB} - \rho_{AB})$. Then it is easy to see that $\frac{1}{2} \| \rho_{AB} - \sigma_{AB} \|_1 = \varepsilon$ and that

$$I(A: B)_{\rho} - I(A: B)_{\sigma} = H(\sigma_{AB}) - H(\rho_{AB}) = 2\varepsilon \log d + h_2(\varepsilon) + O(\varepsilon/d^2).$$

□

**Remark 1.** By using Audenaert’s continuity bound (3) and Winter’s continuity bound (6) one can obtain via representation (18) with trivial $C$ the following continuity bound

$$|I(A: B)_{\rho} - I(A: B)_{\sigma}| \leq \varepsilon \log(d - 1) + 2\varepsilon \log d + h_2(\varepsilon) + g(\varepsilon),$$

□
for the quantum mutual information (for \( \varepsilon \leq 1 - 1/d \)). Since \( h_2(\varepsilon) \leq g(\varepsilon) \) for \( \varepsilon > 0 \), this continuity bound is slightly better than (20) for \( d = 2 \).

Consider the states

\[
\rho_{ABC} = \sum_{i=1}^{m} p_i \rho^i_{AC} \otimes |i\rangle\langle i| \quad \text{and} \quad \sigma_{ABC} = \sum_{i=1}^{m} q_i \sigma^i_{AC} \otimes |i\rangle\langle i|, \tag{21}
\]

where \( \{p_i, \rho^i_{AC}\}_{i=1}^{m} \) and \( \{q_i, \sigma^i_{AC}\}_{i=1}^{m} \) are ensemble of \( m \leq +\infty \) quantum states in \( \mathcal{S}(H_{AC}) \) and \( \{|i\rangle\}_{i=1}^{m} \) is an orthonormal basis in \( H_B \). Such states are called qqc-states in [31]. It follows from upper bound (9) that

\[
I(A:B|C)_\rho \leq I(AC:B)_\rho \leq \max\{H(\rho_{AC}), H(\rho_B)\} \tag{22}
\]

for any qqc-state \( \rho_{ABC} \).

**Corollary 2.** If \( \rho_{ABC} \) and \( \sigma_{ABC} \) are qqc-states (21) then

\[
|I(A:B|C)_\rho - I(A:B|C)_\sigma| \leq \varepsilon \log d + 2g(\varepsilon), \tag{23}
\]

where \( d \doteq \min\{\dim H_{AC}, m\} \) and \( \varepsilon = \frac{1}{2}\|\rho - \sigma\|_1 \).

The first term in (23) can be replaced by \( \varepsilon \max\{S(\{\gamma_i^-\}), S(\{\gamma_i^+\})\} \), where \( \gamma_i^\pm = (2\varepsilon)^{-1}(\|p_i \rho^i_{AC} - q_i \sigma^i_{AC}\|_1 \pm (p_i - q_i)) \) and \( S \) is the Shannon entropy.

**Proof.** The both assertions follow from Proposition 1 and upper bound (22), since

\[
\tau_\pm = \frac{1}{\varepsilon} \sum_{i=1}^{m} [p_i \rho^i_{AC} - q_i \sigma^i_{AC}] \pm \otimes |i\rangle\langle i| \quad \text{and hence} \quad [\tau_\pm]_B = \sum_{i=1}^{m} \gamma_i^\pm |i\rangle\langle i|. \square
\]

If \( \rho_{ABC} \) is a qqc-state (21) then it is easy to show that

\[
I(A:B|C)_\rho = \chi(\{p_i, \rho^i_{AC}\}) - \chi(\{p_i, \rho^i_C\}),
\]

where \( \chi(\{p_i, \rho^i\}) \) is the Holevo quantity of ensemble \( \{p_i, \rho^i\} \). So, Corollary 2 with trivial \( C \) gives continuity bound for the Holevo quantity as a function of ensemble (see Section 4). Corollary 2 with nontrivial \( C \) can be used in analysis of the loss of the Holevo quantity under action of a quantum channel.
3.2 Winter’s type continuity bound for $I(A:B|C)$.

If the both systems $A$ and $B$ are infinite-dimensional (and $C$ is arbitrary) then the function $I(A:B|C)_\omega$ is not continuous on $\mathcal{S}(\mathcal{H}_{ABC})$ (only lower semicontinuous) and takes infinite values. Several conditions of local continuity of this function are presented in Corollary 7 in [28], which implies, in particular, that the function $I(A:B|C)_\omega$ is continuous on subsets of tripartite states $\omega_{ABC}$ with bounded energy of $\omega_A$, i.e. subsets of the form

$$\mathcal{S}_E \doteq \{ \omega_{ABC} | \text{Tr} H_A \omega_A \leq E \},$$

where $H_A$ is the Hamiltonian of system $A$ and $E > 0$, provided that\(^5\)

$$\text{Tr} e^{-\beta H_A} < +\infty \text{ for all } \beta > 0. \quad (25)$$

Condition (25) implies that $H_A$ has discrete spectrum of finite multiplicity, i.e. $H_A = \sum_{n=0}^{+\infty} E_n |n\rangle \langle n|$, where $\{|n\rangle\}_{n=0}^{+\infty}$ is an orthonormal basis of eigenvectors of $H_A$ corresponding to the nondecreasing sequence $\{E_n\}_{n=0}^{+\infty}$ of eigenvalues (energy levels of $H_A$) such that $\sum_{n=0}^{+\infty} e^{-\beta E_n}$ is finite for all $\beta > 0$. We will assume for simplicity that

$$E_0 = \inf_{\|\varphi\|=1} \langle \varphi | H_A | \varphi \rangle = 0. \quad (26)$$

By condition (25) for any $E > 0$ the von Neumann entropy $H(\rho)$ attains its unique maximum under the constraint $\text{Tr} H_A \rho \leq E$ at the Gibbs state $\gamma_A(E) = [\text{Tr} e^{-\beta(E) H_A}]^{-1} e^{-\beta(E) H_A}$, where $\beta(E)$ is the solution of the equation $\text{Tr} H_A e^{-\beta H_A} = E \text{Tr} e^{-\beta H_A}$ [30].

Winter’s type tight continuity bound for the function $I(A:B|C)_\omega$ on the subset $\mathcal{S}_E$ is presented in [29, the Appendix]. The following proposition contains refinement of this bound obtained by using Corollary [11]

**Proposition 2.** Let $H_A$ be the Hamiltonian of system $A$ satisfying conditions (23) and (26). Let $\rho$ and $\sigma$ be any states in $\mathcal{S}(\mathcal{H}_{ABC})$ such that $\text{Tr} H_{A\rho A}, \text{Tr} H_A \sigma_A \leq E$, $\frac{1}{2} \| \rho - \sigma \|_1 \leq \varepsilon < \varepsilon' \leq 1$ and $\delta = \frac{\varepsilon' - \varepsilon}{1 + \varepsilon'}$. Then

$$|I(A:B|C)_\rho - I(A:B|C)_\sigma| \leq (2\varepsilon' + 4\delta) H(\gamma_A(E/\delta)) + 2g(\varepsilon') + 4h_2(\delta), \quad (27)$$

where $g(x) = (1 + x) h_2 \left( \frac{x}{1 + x} \right)$. Continuity bound (27) is asymptotically tight for large $E$ even for trivial $C$, i.e. in the case $I(A:B|C) = I(A:B)$.\(^6\)

\(^5\)Since condition (25) guarantees continuity of the entropy $H(\omega_A)$ on the set $\mathcal{S}_E$ [30].

\(^6\)We say that a continuity bound $|f(x) - f(y)| \leq B(x,y)$ depending on a parameter $a$ is asymptotically tight for large $a$ if $\limsup_{a \to +\infty} \sup_{x,y} \frac{|f(x) - f(y)|}{B(x,y)} = 1$. 

11
Remark 2. A freedom of choice of $\varepsilon'$ makes continuity bound (27) more effective (see [32], where similar continuity bounds for the entropy and for the conditional entropy are obtained).

Remark 3. Condition (25) implies $\lim_{\delta \to +0} \delta H(\gamma_A(E/\delta)) = 0$ [27, Pr.1]. Hence, Proposition 2 shows that the function $\omega_{ABC} \mapsto I(A:B|C)_\omega$ is uniformly continuous on the set $\mathcal{S}_E$ for any $E > 0$ (one can take $\varepsilon' = \sqrt{\varepsilon}$).

Proof. The proof of continuity bound (27) differs from the proof of Lemma 25 in [29] only by using Corollary 1 instead of Corollary 8 in [28]. The asymptotic tightness of continuity bound (27) follows from the asymptotic tightness of the continuity bound in Corollary 3 (see Remark 5 below). □

Assume now that $A$ is the system composed of $\ell$ quantum oscillators (while $B$ and $C$ are arbitrary systems). The Hamiltonian of such system has the form

$$H_A = \sum_{i=1}^{\ell} \hbar \omega_i a_i^+ a_i,$$

where $a_i$ and $a_i^+$ are the annihilation and creation operators and $\omega_i$ is a frequency of the $i$-th oscillator [13]. To be consistent with our assumption $E_0 = 0$ we will consider shifted Hamiltonian $H'_A = H_A - \frac{1}{2} \sum_{i=1}^{\ell} \hbar \omega_i I_A.$

By repeating the arguments from the proof of Lemma 18 in [32] with Proposition 2 instead of Meta-Lemmas 16,17 one can obtain the following

**Corollary 3.** Let $A$ be the system of $\ell$ quantum oscillators. Let $\rho$ and $\sigma$ be any states in $\mathcal{S}(H_{ABC})$ such that $\text{Tr} H'_A \rho_A, \text{Tr} H'_A \sigma_A \leq E$ and $\frac{1}{2} \|\rho - \sigma\|_1 \leq \varepsilon$. Then

$$|I(A:B|C)_\rho - I(A:B|C)_\sigma| \leq 2\varepsilon \left(\frac{1+\alpha}{1-\alpha} + 2\alpha\right) \left[ \sum_{i=1}^{\ell} \log \left( \frac{E}{\hbar \omega_i} + 1 \right) + \ell \log \frac{e}{\alpha(1-\varepsilon)} \right]$$

$$+ 2\ell \left(\frac{1+\alpha}{1-\alpha} + 2\alpha\right) \tilde{h}_2(\varepsilon) + 4\tilde{h}_2(\alpha\varepsilon) + 2g \left(\frac{1+\alpha}{1-\alpha} \varepsilon\right),$$

where $\alpha \in (0, \frac{1}{2})$, $\tilde{h}_2(x) = h_2(x)$ for $x \leq 1/2$ and $\tilde{h}_2(x) = 1$ for $x \geq 1/2$, $g(x) = (x + 1) \log(x + 1) - x \log x$.

7This means that the energy of $\rho$ is equal to $\text{Tr} H'_A \rho + \frac{1}{2} \sum_{i=1}^{\ell} \hbar \omega_i$. 

12
Remark 4. Parameter \( \alpha \) in Corollary 3 is a free parameter which can be used to optimize the continuity bound for given value of energy \( E \). The below Lemma 1 (proved by elementary methods) implies that for large energy \( E \) the main term in this continuity bound can be made not greater than \( \varepsilon (2A_E + o(A_E)) \) by appropriate choice of \( \alpha \), where

\[
A_E = \sum_{i=1}^{\ell} \log \left( \frac{E}{\ell \omega_i} + 1 \right) \approx H(\gamma_A(E)).
\]

Lemma 1. Let \( f(\alpha) = \frac{a + \alpha}{1 - \alpha} + 2\alpha, \ a > 0 \) and \( b \) be arbitrary. Then

\[
\min_{\alpha \in (0, \frac{1}{2})} f(\alpha)(x - a \log \alpha + b) \leq x + o(x), \ x \to +\infty.
\]

Remark 5. Remark 4 makes it possible to show the asymptotic tightness of the continuity bound in Corollary 3 for trivial \( C \). Indeed, let \( \rho \) be a purification of the Gibbs state \( \gamma_A(E) \) and \( \sigma = (1 - \varepsilon)\rho + \varepsilon \alpha \otimes \beta \), where \( \alpha \) is a state in \( \mathcal{S}(H_A) \) such that \( \text{Tr} H_A\alpha \leq E \) and \( \beta \) is any state in \( \mathcal{S}(H_B) \). Then inequality (17) implies

\[
I(A:B)_{\rho} - I(A:B)_{\sigma} \geq 2\varepsilon H(\gamma_A(E)) - h_2(\varepsilon).
\]

3.3 Continuity bound for the function \( \Phi \mapsto I(B^n : D|C)_{\Phi^\otimes n}^{\otimes n}(\rho) \)

The following proposition is a CMI-analog of Theorem 11 in [18] proved by the same telescopic trick. It gives Fannes’ type and Winter’s type tight continuity bounds for the function \( \Phi \mapsto I(B^n : D|C)_{\Phi^\otimes n}^{\otimes n}(\rho) \) for any given \( n \) and a state \( \rho \in \mathcal{S}(H_A^\otimes n \otimes H_{CD}) \) with respect to the diamond norm on the set of all channels from \( A \) to \( B \) (described at the end of Section 2).

Proposition 3. Let \( \Phi \) and \( \Psi \) be channels from \( A \) to \( B \), \( \varepsilon = \frac{1}{2}\|\Phi - \Psi\|_\diamond \), \( C \) and \( D \) be any systems. Let \( \rho \) be any state in \( \mathcal{S}(H_A^\otimes n \otimes H_{CD}) \), \( n \in \mathbb{N} \), and

\[
\Delta^n(\Phi, \Psi, \rho) = \left| I(B^n : D|C)_{\Phi^\otimes n \otimes \text{Id}_{CD}(\rho)}^{\otimes n} - I(B^n : D|C)_{\Psi^\otimes n \otimes \text{Id}_{CD}(\rho)}^{\otimes n} \right|.
\]

A) If \( d_B = \dim H_B < +\infty \) then

\[
\Delta^n(\Phi, \Psi, \rho) \leq 2n\varepsilon \log d_B + 2ng(\varepsilon). \tag{28}
\]
where it was used that $\text{Tr}_{B_k} \sigma_k = \text{Tr}_{B_k} \sigma_{k-1}$. By upper bound (16) the finiteness of the entropy of the states $[\sigma_k]_{B_1}, ..., [\sigma_k]_{B_n}$ guarantees finiteness of all the terms in (30) and (31).
Since \( \| \sigma_k - \sigma_{k-1} \|_1 \leq \| \Phi - \Psi \|_\infty = 2 \varepsilon \), by applying Corollary 1 to the right hand side of (31) in case A we obtain that the value

\[
\left| I(B^n : D | C)_{\sigma_k} - I(B^n : D | C)_{\sigma_{k-1}} \right|
\]

is upper bounded by \( 2 \varepsilon \log d_B + 2 g(\varepsilon) \) for any \( k \). Similarly, by using Proposition 2 in case B we obtain that for any \( k \) the value (32) is upper bounded by \( (2\varepsilon' + 4\delta)H(\gamma_B(E_k/\delta)) + 2g(\varepsilon') + 4h_2(\delta) \). Hence (28) and the first inequality in (29) follow from (30) (since \( \Phi^\otimes n \otimes \text{Id}_{CD}(\rho) = \sigma_n \) and \( \Psi^\otimes n \otimes \text{Id}_{CD}(\rho) = \sigma_0 \)). The second inequality in (29) follows from the concavity of the function \( E \mapsto H(\gamma_B(E)) \) Proposition 11.

The tightness of the continuity bound (28) for trivial \( C \) and any given \( n \) can be shown by using the erasure channels

\[
\Phi_p(\rho) = \begin{bmatrix}
(1 - p)\rho & 0 \\
0 & p \text{Tr} \rho
\end{bmatrix}, \quad p \in [0, 1].
\]

from \( d \)-dimensional system \( A \) to \( d + 1 \)-dimensional system \( B \). Indeed, let \( D \cong A \) and \( \rho \) be any maximally entangled pure state in \( \mathcal{S}(\mathcal{H}_{AD}) \). Then \( I(B : D)_{\Phi_0 \otimes \text{Id}_D(\rho)} = 2 \log d_A \) and by using inequality (17) it is easy to show that \( I(B : D)_{\Phi_p \otimes \text{Id}_D(\rho)} \leq 2(1 - p) \log d_A + h_2(p) \). So, we have

\[
I(B^n : D^n)_{\Phi_0^\otimes n \otimes \text{Id}_{D^n}(\rho^\otimes n)} = nI(B : D)_{\Phi_0 \otimes \text{Id}_D(\rho)} = 2n \log d_A,
\]

\[
I(B^n : D^n)_{\Phi_p^\otimes n \otimes \text{Id}_{D^n}(\rho^\otimes n)} = nI(B : D)_{\Phi_p \otimes \text{Id}_D(\rho)} \leq 2n(1 - p) \log d_A + nh_2(p)
\]

and hence

\[
I(B^n : D^n)_{\Phi_0^\otimes n \otimes \text{Id}_{D^n}(\rho^\otimes n)} - I(B^n : D^n)_{\Phi_p^\otimes n \otimes \text{Id}_{D^n}(\rho^\otimes n)} \geq 2np \log d_A - nh_2(p)
\]

Since \( d_B = d_A + 1 \) and \( \| \Phi_0 - \Phi_p \|_\infty \leq 2p \), this shows the tightness of the continuity bound (28) for large \( d_B \).

The asymptotic tightness of the continuity bound (29) for trivial \( C \) and any given \( n \) can be shown by using the erasure channels (33) from the system \( A \) composed of \( \ell \) quantum oscillators to any its one-dimensional extension \( B \). If \( \rho \) is any purification of the Gibbs state \( \gamma(E) \) then the above arguments imply

\[
I(B^n : D^n)_{\Phi_0^\otimes n \otimes \text{Id}_{D^n}(\rho^\otimes n)} - I(B^n : D^n)_{\Phi_p^\otimes n \otimes \text{Id}_{D^n}(\rho^\otimes n)} \geq 2npH(\gamma(E)) - nh_2(p)
\]

This shows the asymptotic tightness of the continuity bound (29) for large \( E \), since in this case the main term of (29) can be made not greater than
\[ \epsilon [2H(\gamma(E)) + o(H(\gamma(E)))], \] for large \( E \) by appropriate choice of \( \epsilon' \) (see Remark 4 in Section 3.2). \( \square \)

By using Proposition 3A one can obtain tight and close-to-tight continuity bounds for quantum and classical capacities of finite-dimensional channels (essentially refining the Leung-Smith continuity bounds), Proposition 3B makes it possible to obtain close-to-tight continuity bound for the classical capacity of infinite-dimensional quantum channels with finite energy amplification factors (see Sections 5.2 and 5.3 below).

4 On continuity of the Holevo quantity

The Holevo quantity of an ensemble \( \{p_i, \rho_i\}_{i=1}^m \) of \( m \leq +\infty \) quantum states is defined as

\[
\chi (\{p_i, \rho_i\}_{i=1}^m) = \sum_{i=1}^m p_i H(\rho_i || \bar{\rho}) = H(\bar{\rho}) - \sum_{i=1}^m p_i H(\rho_i), \quad \bar{\rho} = \sum_{i=1}^m p_i \rho_i,
\]

where the second formula is valid if \( H(\bar{\rho}) < +\infty \). This quantity gives the upper bound for classical information which can be obtained by applying quantum measurements to an ensemble [12]. It plays important role in analysis of information properties of quantum systems and channels [13, 23, 31].

Let \( \mathcal{H}_A = \mathcal{H} \) and \( \{|i\}_{i=1}^m \) be an orthonormal basis in a \( m \)-dimensional Hilbert space \( \mathcal{H}_B \). Then

\[
\chi(\{p_i, \rho_i\}_{i=1}^m) = I(A:B)_{\hat{\omega}}, \text{ where } \hat{\omega}_{AB} = \sum_{i=1}^m p_i \rho_i \otimes |i\rangle \langle i|.
\] (34)

If \( H(\bar{\rho}) \) and \( S(\{p_i\}_{i=1}^m) \) are finite (here \( S \) is the Shannon entropy) then (34) is directly verified by noting that \( H(\hat{\omega}_A) = H(\bar{\rho}), \; H(\hat{\omega}_B) = S(\{p_i\}_{i=1}^m) \) and \( H(\hat{\omega}_{AB}) = \sum_{i=1}^m p_i H(\rho_i) + S(\{p_i\}_{i=1}^m) \). The validity of (34) in general case can be easily shown by two step approximation using Theorem 1A in [28].

To analyse continuity of the Holevo quantity as a function of an ensemble we have to choose a measure of divergence between ensembles.
4.1 Two nonequivalent metrics on the set of quantum ensembles

If we consider an ensemble as an ordered collection of states with the corresponding probability distribution then it is natural to use the quantity

\[ D_0(\mu, \nu) = \frac{1}{2} \sum_i \| p_i \rho_i - q_i \sigma_i \|_1 \]

as a distance between ensembles \( \mu = \{ p_i, \rho_i \} \) and \( \nu = \{ q_i, \sigma_i \} \). Since \( D_0(\mu, \nu) \) coincides (up to the factor \( 1/2 \)) with the trace norm of the difference between the corresponding cq-states \( \sum_i p_i \rho_i \otimes |i\rangle \langle i| \) and \( \sum_i q_i \sigma_i \otimes |i\rangle \langle i| \), \( D_0 \) is a true metric on the set of all "ordered" ensembles of quantum states. Since convergence of a sequence of states to a state in the weak operator topology implies convergence of this sequence in the trace norm [8], a sequence \( \{ p^n_i, \rho^n_i \}_n \) of ensembles converges to an ensemble \( \{ p^0_i, \rho^0_i \} \) with respect to the metric \( D_0 \) if and only if

\[ \lim_{n \to \infty} p^n_i = p^0_i \quad \text{for all } i \quad \text{and} \quad \lim_{n \to \infty} \rho^n_i = \rho^0_i \quad \text{for all } i \text{ such that } p^0_i \neq 0. \]

(35)

But from the quantum information point of view (in particular, in analysis of the Holevo quantity) it is reasonable to consider an ensemble of quantum states \( \{ p_i, \rho_i \} \) as a discrete probability measure \( \sum_i p_i \delta(\rho_i) \) on the set \( \mathcal{G}(\mathcal{H}) \) (where \( \delta(\rho) \) is the Dirac measure concentrating at a state \( \rho \)) rather than ordered (or disordered) collection of states. It suffices to say that a singleton ensemble consisting of a state \( \sigma \) and the ensemble \( \{ p_i, \rho_i \} \), where \( \rho_i = \sigma \) for all \( i \), are identical from the information point of view and correspond to the same measure \( \delta(\sigma) \).

For any ensemble \( \{ p_i, \rho_i \} \) denote by \( E(\{ p_i, \rho_i \}) \) the set of all countable ensembles corresponding to the measure \( \sum_i p_i \delta(\rho_i) \). The set \( E(\{ p_i, \rho_i \}) \) consists of ensembles obtained from the ensemble \( \{ p_i, \rho_i \} \) by composition of the following operations:

- permutation of any states;
- splitting: \( (p_1, \rho_1), (p_2, \rho_2), ... \to (p, \rho_1), (p_1-p, \rho_1), (p_2, \rho_2), ..., p \in [0, p_1] \);
- joining of equal states: \( (p_1, \rho_1), (p_2, \rho_1), (p_3, \rho_3), ... \to (p_1+p_2, \rho_1), (p_3, \rho_3), ... \)
If we want to identify ensembles corresponding to the same probability measure then it is natural to use the factorization of $D_0$, i.e. the quantity

$$D_*(\mu, \nu) \doteq \inf_{\mu' \in E(\mu), \nu' \in E(\nu)} D_0(\mu', \nu')$$

as a measure of divergence between ensembles $\mu$ and $\nu$.

The problem of finding appropriate "distinguishability measures" between ensembles of quantum states is considered by Oreshkov and Calsamiglia in [24]. In particular, they proposed to use in the role of such measure the EHS-distance

$$D_{\text{ehs}}(\mu, \nu) = \frac{1}{2} \inf_{P, Q} \sum_{i,j} \| P_{ij} \rho_i - Q_{ij} \sigma_j \|_1$$

between ensembles $\mu = \{p_i, \rho_i\}$ and $\nu = \{q_i, \sigma_i\}$, where the infimum is over all joint probability distributions $P \doteq \{P_{ij}\}$ with the left marginal $\{p_i\}$ and $Q \doteq \{Q_{ij}\}$ with the right marginal $\{q_j\}$. It is shown in [24] that $D_{\text{ehs}}$ is a true metric on the sets of discrete ensembles (considered as probability measures) having operational interpretations and possessing several natural properties (convexity, monotonicity under action of quantum channels and generalized measurements, etc.).

The following proposition is proved in the Appendix.

**Proposition 4.** A) The factor-metric $D_*$ and the metric $D_{\text{ehs}}$ (defined respectively by (36) and (37)) coincide on the set of all discrete ensembles.

B) The metric $D_* = D_{\text{ehs}}$ generates the weak convergence topology on the set of all ensembles (considered as probability measures), i.e. convergence of a sequence $\{\{p^n_i, \rho^n_i\}\}_n$ to an ensemble $\{p^0_i, \rho^0_i\}$ with respect to the metric $D_* = D_{\text{ehs}}$ means that

$$\lim_{n \to \infty} \sum_i p^n_i f(\rho^n_i) = \sum_i p^0_i f(\rho^0_i)$$

for any continuous bounded function $f$ on $\mathcal{S}(\mathcal{H})$.

**Remark 6.** The coincidence of $D_*$ and $D_{\text{ehs}}$ shows, in particular, that for ensembles $\mu$ and $\nu$ consisting of $m$ and $n$ states correspondingly the infimum

---

8The abbreviation "EHS" means "Extended Hilbert Space", it is justified by the fact that $D_{\text{ehs}}(\mu, \nu)$ is (up to the factor 1/2) the infimum of the trace norm distance between the $eq$-states $\sum_{i,j} P_{ij} \rho_i \otimes |i\rangle \langle i| \otimes |j\rangle \langle j|$ and $\sum_{i,j} Q_{ij} \sigma_j \otimes |i\rangle \langle i| \otimes |j\rangle \langle j|$ [24].
in (36) is attained at some ensembles $\mu'$ and $\nu'$ consisting of $\leq mn$ states and that it can be calculated by standard linear programming procedure [24].

The weak convergence topology is widely used in the measure theory and its applications [7, 25]. It has different characterizations. In particular, Theorem 6.1 in [25] shows that the weak convergence of a sequence $\left\{ \{ p^i_n, \rho^i_n \} \right\}_n$ to an ensemble $\{ p^0_i, \rho^0_i \}$ means that

$$\lim_{n \to \infty} \sum_{i: \rho^i_n \in \mathcal{G}} p^i_n = \sum_{i: \rho^0_i \in \mathcal{G}} p^0_i$$

for any subset $\mathcal{G}$ of $\mathcal{S}(H)$ such that $\{ \rho^0_i \} \cap \partial \mathcal{G} = \emptyset$ ($\partial \mathcal{G}$ is the boundary of $\mathcal{G}$). It is easy to see that this convergence is substantially weaker than convergence (35).

Despite the fact that the metric $D_* = D_{ehs}$ is more adequate for analysis of the Holevo quantity, the metric $D_0$ will be also used in what follows. The main advantage of $D_0$ is its simple computability. Moreover, in some cases the metrics $D_0$ and $D_* = D_{ehs}$ is close to each other or even coincide. This holds, for example, if we consider small perturbations of states or probabilities of a given ensemble.

So, we will explore continuity of the function $\{ p_i, \rho_i \} \mapsto \chi(\{ p_i, \rho_i \})$ with respect to both metrics $D_0$ and $D_* = D_{ehs}$, i.e. with respect to the convergence (35) and to the weak convergence (38). We will obtain Fannes’ type and Winter’s type continuity bounds for this function with respect to the above two metrics.

### 4.2 The case of global continuity

The following proposition contains continuity bounds for the Holevo quantity with respect to the metrics $D_0$ and $D_* = D_{ehs}$ (denoted $D_*$ in what follows).

**Proposition 5.** Let $\{ p_i, \rho_i \}$ and $\{ q_i, \sigma_i \}$ be arbitrary ensembles of states in $\mathcal{S}(H)$, $\varepsilon_0 = D_0(\{ p_i, \rho_i \}, \{ q_i, \sigma_i \})$ and $\varepsilon_* = D_*(\{ p_i, \rho_i \}, \{ q_i, \sigma_i \})$.

A) If $d = \dim H$ is finite then

$$|\chi(\{ p_i, \rho_i \}) - \chi(\{ q_i, \sigma_i \})| \leq \varepsilon_* \log d + 2g(\varepsilon_*) \leq \varepsilon_0 \log d + 2g(\varepsilon_0),$$

where $g(x) = (1 + x)h_2(\frac{x}{1+x})$.

B) If $\{ p_i, \rho_i \}$ and $\{ q_i, \sigma_i \}$ are ensembles consisting of $m$ and $n \leq m$ states respectively then

$$|\chi(\{ p_i, \rho_i \}) - \chi(\{ q_i, \sigma_i \})| \leq \min\{ \varepsilon_* \log(mn) + 2g(\varepsilon_*), \varepsilon_0 \log m + 2g(\varepsilon_0) \}.$$
The term \( \log m \) in (41) can be replaced by \( \max\{S(\{\gamma^-\}), S(\{\gamma^+\})\} \), where \( \gamma_i^\pm = (2\varepsilon_0)^{-1}(\|p_i\rho_i - q_i\sigma_i\|_1 \pm (p_i - q_i)), i = 1, m \), \( S \) is the Shannon entropy and it is assumed that \( q_i = 0 \) for \( i > n \) (if \( n < m \)).

The both continuity bounds in (40) and the both continuity bounds in (41) are tight.

Proof. The second continuity bounds in (40) and in (41) and the specification of the latter follow from representation (34) and Corollary 2 with trivial \( C \).

Take any joint probability distributions \( P = \{P_{ij}\} \) with the left marginal \( \{p_i\} \) and \( Q = \{Q_{ij}\} \) with the right marginal \( \{q_j\} \) and consider the qc-states

\[
\hat{\rho}_{ABC} = \sum_{i,j} P_{ij} p_i |i\rangle \langle i| \otimes |j\rangle \langle j|, \quad \hat{\sigma}_{ABC} = \sum_{i,j} Q_{ij} q_j |i\rangle \langle i| \otimes |j\rangle \langle j|, \quad (42)
\]

where \( \{|i\rangle\}_{i=1}^n \) and \( \{|j\rangle\}_{j=1}^n \) are orthonormal base of Hilbert spaces \( \mathcal{H}_B \) and \( \mathcal{H}_C \) correspondingly. Representation (34) and the invariance of the Holevo quantity under splitting of states of an ensemble imply

\[
\chi(\{p_i, \rho_i\}) = I(A:BC)_{\hat{\rho}} \quad \text{and} \quad \chi(\{q_j, \sigma_j\}) = I(A:BC)_{\hat{\sigma}}. \quad (43)
\]

Thus, the first continuity bounds in (40) and in (41) also follow from Corollary 2 with trivial \( C \) (since \( 2\varepsilon_* = \inf\|\hat{\rho} - \hat{\sigma}\|_1 \), where the infimum is over all states (42)).

Let \( \{|i\rangle\}_{i=1}^d \) be an orthonormal basis in \( \mathcal{H} = \mathbb{C}^d \) and \( \rho_c = I/d \) the chaotic state in \( \mathcal{G}(\mathcal{H}) \). For given \( \varepsilon \in (0, 1) \) consider the ensembles \( \mu = \{p_i, \rho_i\}_{i=1}^d \) and \( \nu = \{q_i, \sigma_i\}_{i=1}^d \), where \( p_i = \varepsilon |i\rangle \langle i|, \sigma_i = (1 - \varepsilon)|i\rangle \langle i| + \varepsilon \rho_c, p_i = q_i = 1/d \) for all \( i \). Then it is easy to see that \( D_\varepsilon(\mu, \nu) \leq D_0(\mu, \nu) = \varepsilon(1 - 1/d) \), while concavity of the entropy implies

\[
\chi(\mu) - \chi(\nu) = \log d - \log d + H(\sigma_i) \geq \varepsilon \log d.
\]

Since \( \dim \mathcal{H} = m = n = d \), this shows tightness of the both continuity bounds in (40) and of the second continuity bound in (41). This example with \( d = 3 \) also shows that the second terms in (40) can not be less than \( \varepsilon \log 3/3 \approx 0.53\varepsilon \).

Modifying the above example consider the ensemble \( \mu = \{p_i, \rho_i\}_{i=1}^d \), where \( \rho_i = \varepsilon |i\rangle \langle i| + (1 - \varepsilon)\rho_c \) and \( p_i = 1/d \) for all \( i \), and the singleton ensemble
$\nu = \{\rho_i\}$. Then it is easy to see that $D_*(\mu, \nu) \leq \varepsilon$, while inequality \[2\] implies
\[
\chi(\mu) - \chi(\nu) = \chi(\mu) \geq \varepsilon \log d - h_2(\varepsilon).
\]
Since $\dim \mathcal{H} = mn = d$, this shows the tightness of the first continuity bounds in (40) and in (41). Since $D_0(\mu, \nu) \geq (d-1)/d$ for any $\varepsilon$, this example also shows the difference between the continuity bounds depending on $D_*(\mu, \nu)$ and on $D_0(\mu, \nu)$. □

Let $\mathcal{E}_m^0(\mathcal{H})$ and $\mathcal{E}_m^*(\mathcal{H})$ be the sets of all ensembles consisting of $\leq m$ different states equipped with the metric $D_0$ and $D_*$ respectively. By Proposition 4B the set $\mathcal{E}_m^*(\mathcal{H})$ can be treated as the set of discrete probability measures on $\mathcal{G}(\mathcal{H})$ with $\leq m$ atoms equipped with weak convergence topology. Proposition 5 implies Corollary 4.

The function $\{p_i, \rho_i\} \mapsto \chi(\{p_i, \rho_i\})$ is uniformly continuous on $\mathcal{E}_m^*(\mathcal{H})$ if either $\dim \mathcal{H}$ or $m$ is finite. Otherwise this function is not continuous on $\mathcal{E}_m^*(\mathcal{H})$.

Proof. It suffices to show that the function $\{p_i, \rho_i\} \mapsto \chi(\{p_i, \rho_i\})$ is not continuous on $\mathcal{E}_m^0(\mathcal{H})$ if $\dim \mathcal{H} = m = +\infty$.

Let $\{\{\pi_i^0\}_i\}_n$ be a sequence of countable probability distributions converging (in the $\ell_1$-metric) to a probability distribution $\{\pi_i^0\}_i$, such that $S(\{\pi_i^0\}_i) \rightarrow S(\{\pi_i^0\}_i)$ (where $S$ is the Shannon entropy). Let $\{\rho_i\}$ be a countable collection of mutually orthogonal pure states in a separable Hilbert space $\mathcal{H}$. Then the sequence of ensembles $\{\{\pi_i^0, \rho_i\}_i\}_n$ converges to the ensemble $\{\pi_i^0, \rho_i\}_i$ in the sense (35), but $\chi(\{\pi_i^0, \rho_i\}) = S(\{\pi_i^0\}_i)$ do not converge to $\chi(\{\pi_i^0, \rho_i\}) = S(\{\pi_i^0\}_i)$. □

Proposition 5 contains estimates of two types: the continuity bounds with the main term $\varepsilon \log \dim \mathcal{H}$ depending only on the dimension of underlying Hilbert space $\mathcal{H}$ and the continuity bounds with the main term $\varepsilon \log m$ depending only on the size $m$ of ensembles. Continuity bounds of the last type are sometimes called dimension-independent. Recently Audenaert obtained the following dimension-independent continuity bound for the Holevo quantity in the case $p_i \equiv q_i$ [5, Th.15]:
\[
|\chi(\{p_i, \rho_i\}) - \chi(\{p_i, \sigma_i\})| \leq t \log(1 + (m-1)/t) + \log(1 + (m-1)t),
\]
where $t = \frac{1}{2} \max_i \|\rho_i - \sigma_i\|_1$ is the maximal distance between corresponding
states of ensembles. Proposition 5B in this case gives
\[ |\chi(\{p_i, \rho_i\}) - \chi(\{p_i, \sigma_i\})| \leq \varepsilon \log m + 2g(\varepsilon), \tag{44} \]
where \( \varepsilon = \frac{1}{2} \sum_i p_i \|\rho_i - \sigma_i\|_1 \) is the average distance between corresponding states of ensembles. Since \( \varepsilon \leq t \) and \( g(x) \) is an increasing function on \([0, 1]\), we may replace \( \varepsilon \) by \( t \) in (44).

The following continuity bound for the Holevo quantity not depending on the size \( m \) of an ensemble is obtained by Oreshkov and Calsamiglia in [24]:
\[ |\chi(\{p_i, \rho_i\}) - \chi(\{q_i, \sigma_i\})| \leq 2\varepsilon_K \log(d - 1) + 2h_2(\varepsilon_K), \quad \varepsilon_K \leq (d - 1)/d, \]
where \( d = \dim \mathcal{H} \) and \( \varepsilon_K \) is the Kantorovich distance between the ensembles \( \{p_i, \rho_i\} \) and \( \{q_i, \sigma_i\} \). Since the EHS-distance is upper bounded by the Kantorovich distance [24, Pr.9], Proposition 3 implies \( \varepsilon_K \geq \varepsilon_* = D_*(\{p_i, \rho_i\}, \{q_i, \sigma_i\}) \). So, Proposition 5A gives stronger continuity bound for the Holevo quantity for \( d > 2 \).

### 4.3 General case

If \( \dim \mathcal{H} = m = +\infty \) then the Holevo quantity is not continuous on \( \mathcal{E}_m^0(\mathcal{H}) \) and on \( \mathcal{E}_m^*(\mathcal{H}) \). By Proposition 2 in [15] it is lower semicontinuous on \( \mathcal{E}_\infty^*(\mathcal{H}) \) and hence on \( \mathcal{E}_\infty^0(\mathcal{H}) \). Conditions for local continuity of the Holevo quantity are presented in the following proposition.

**Proposition 6.** A) If \( \{\{p_i^n, \rho_i^n\}\}_n \) is a sequence of countable ensembles weakly converging to an ensemble \( \{p_i^0, \rho_i^0\} \) and
\[ \lim_{n \to \infty} H \left( \sum_i p_i^n \rho_i^n \right) = H \left( \sum_i p_i^0 \rho_i^0 \right) < +\infty \]
then
\[ \lim_{n \to \infty} \chi(\{p_i^n, \rho_i^n\}) = \chi(\{p_i^0, \rho_i^0\}) < +\infty. \tag{45} \]

B) If \( \{\{p_i^n, \rho_i^n\}\}_n \) is a sequence converging to an ensemble \( \{p_i^0, \rho_i^0\} \) in the sense (35) and
\[ \lim_{n \to \infty} S(\{p_i^n\}) = S(\{p_i^0\}) < +\infty, \tag{46} \]
where \( S \) is the Shannon entropy, then (45) holds.
C) If \( \{\{p^n_i, \rho^n_i\}\}_n \) is a sequence converging to an ensemble \( \{p^0_i, \rho^0_i\} \) in the sense (32) and (45) holds then \( \lim_{n \to \infty} \chi(\{p^n_i, \Phi(\rho^n_i)\}) = \chi(\{p^0_i, \Phi(\rho^0_i)\}) \) for arbitrary quantum channel \( \Phi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}') \).

**Remark 7.** By modifying the example from the proof of Corollary 4 one can show that condition (46) does not imply (45) for weakly converging sequence \( \{\{p^n_i, \rho^n_i\}\}_n \).

**Proof.** A) We may assume that \( H(\bar{\rho}_n) < +\infty \) for all \( n \), where \( \bar{\rho}_n = \sum_i p^n_i \rho^n_i \). So, we have

\[
\chi(\{p^n_i, \rho^n_i\}) = H(\bar{\rho}_n) - \sum_i p^n_i H(\rho^n_i).
\]

Since the function \( \{p_i, \rho_i\} \mapsto \chi(\{p^n_i, \rho^n_i\}) \) is lower semicontinuous on \( \mathcal{E}_m^*(\mathcal{H}) \), to prove (45) it suffice to show that the function \( \{p_i, \rho_i\} \mapsto \sum_i p_i H(\rho_i) \) is lower semicontinuous on \( \mathcal{E}_m^*(\mathcal{H}) \). This can be done by representing the von Neumann entropy \( H \) as a limit of an increasing sequence of continuous bounded functions (see the proof of Proposition 2 in [15]).

B,C) Since convergence (35) implies the trace norm convergence of the sequence \( \{\tilde{\omega}^n_{AB}\} \) to the state \( \tilde{\omega}^0_{AB} \), where \( \tilde{\omega}^n_{AB} = \sum_i p^n_i \rho^n_i \otimes |i\rangle \langle i| \), assertions B and C are derived respectively from Theorems 1A and 1B in [28] by means of representation (34). \( \Box \).

Proposition 7B implies the following observation which can be interpreted as stability of the Holevo quantity with respect to perturbation of states of a given ensemble.

**Corollary 5.** Let \( \{p_i\} \) be a probability distribution with finite Shannon entropy. Then

\[
\lim_{n \to \infty} \chi(\{p_i, \rho^n_i\}) = \chi(\{p_i, \rho^0_i\}) \leq S(\{p_i\})
\]

for any sequences \( \{\rho^1_i\}, \{\rho^2_i\}, \ldots \) converging respectively to states \( \rho^0_1, \rho^0_2, \ldots \).

By Corollary 5 the finiteness of \( S(\{p_i\}) \) guarantees the validity of (47) even in the case when the entropy is not continuous for all the sequences \( \{\rho^n_1\}, \{\rho^n_2\}, \ldots \), i.e. when \( H(\rho^k) \nrightarrow H(\rho^0_k) \) for all \( k = 1, 2, \ldots \).

Proposition 7A shows that for any \( E > 0 \) the Holevo quantity is continuous on the subset of \( \mathcal{E}_\infty^*(\mathcal{H}) \) consisting of ensembles \( \{p_i, \rho_i\} \) with the mean energy \( \text{Tr} H \rho \leq E \) provided the Hamiltonian \( H \) satisfies condition (25).
The following proposition gives Winter’s type continuity bound for the Holevo quantity with respect to the metric $D_*$ under the mean energy constraint.

**Proposition 7.** Let $H_A$ be the Hamiltonian of system $A$ satisfying conditions (25) and (26). Let $\{p_i, \rho_i\}$ and $\{q_i, \sigma_i\}$ be countable ensembles of states in $\mathcal{S}(H_A)$ with the average states $\bar{\rho}$ and $\bar{\sigma}$ such that $\text{Tr}H_A\bar{\rho}, \text{Tr}H_A\bar{\sigma} \leq E$, $D_*(\{p_i, \rho_i\}, \{q_i, \sigma_i\}) = \varepsilon < \varepsilon' \leq 1$ and $\delta = \frac{\varepsilon - \varepsilon'}{1 + \varepsilon'}$. Then

$$|\chi(\{p_i, \rho_i\}) - \chi(\{q_i, \sigma_i\})| \leq (\varepsilon' + 2\delta)H(\gamma_A(E/\delta)) + 2g(\varepsilon') + 2h_2(\delta),$$

where $g(\varepsilon) = (1 + \varepsilon)h_2(\frac{\varepsilon}{1+\varepsilon})$ and $\gamma_A(E)$ is the Gibbs state corresponding to the energy $E$. This continuity bound is asymptotically tight for large $E$.

**Remark 8.** Condition (25) implies $\lim_{\delta \to +0} \delta H(\gamma_A(E/\delta)) = 0$ [27, Pr.1]. Hence, Proposition 7 shows that the Holevo quantity is uniformly continuous with respect to the metric $D_*$ on the set of all ensembles $\{p_i, \rho_i\}$ with bounded mean energy.

**Remark 9.** The metric $D_*$ in Proposition 7 can be replaced by the easy-computable metric $D_0$.

**Proof.** By using representation (43) it is easy to see that continuity bound (48) can be proved by showing that

$$|I(A:B)_\rho - I(A:B)_\sigma| \leq (\varepsilon' + 2\delta)H(\gamma_A(E/\delta)) + 2g(\varepsilon') + 2h_2(\delta)$$

for arbitrary qc-states $\rho_{AB}$ and $\sigma_{AB}$ such that $\text{Tr}H_A\rho_A, \text{Tr}H_A\sigma_A \leq E$ and $\|\rho_{AB} - \sigma_{AB}\|_1 = 2\varepsilon$.

Let $H_A = \sum_{n=0}^{+\infty} E_n |n\rangle\langle n|$. Following the proofs of Lemmas 16,17 in [32] define the projector

$$P_\delta = \sum_{0 \leq E_n \leq E/\delta} |n\rangle\langle n|$$

in $\mathfrak{B}(H_A)$ and consider the states

$$\rho_{AB}^\delta = \frac{P_\delta \otimes I_B \rho_{AB} P_\delta \otimes I_B}{\text{Tr}P_\delta \rho_A}, \quad \sigma_{AB}^\delta = \frac{P_\delta \otimes I_B \sigma_{AB} P_\delta \otimes I_B}{\text{Tr}P_\delta \sigma_A}.$$ 

In the proof of Lemma 16 in [32] it is shown that

$$H(\omega_A) - [\text{Tr}P_\delta \omega_A]H(\omega_A^\delta) \leq \delta H(\gamma_A(E/\delta)) + h_2(\text{Tr}P_\delta \omega_A),$$

(50)
\[ H(\omega_A^\delta) \leq H(\gamma_A(E/\delta)), \quad \text{Tr} P_\delta \omega_A \geq 1 - \delta, \]  
(51)

where \( \omega = \rho, \sigma \), and that
\[ \log \text{Tr} P_\delta \leq H(\gamma_A(E/\delta)), \quad \frac{1}{2} \| \rho^\delta_{AB} - \sigma^\delta_{AB} \|_1 \leq \varepsilon'. \]  
(52)

By using (50) and (51) it is easy to derive from Lemma 2 below that
\[ |I(A:B)_\omega - I(A:B)_{\omega^\delta}| \leq \delta H(\gamma_A(E/\delta)) + h_2(\delta), \quad \omega = \rho, \sigma. \]  
(53)

By using (52) and applying Corollary 2 with trivial \( C \) we obtain
\[ |I(A:B)_{\rho^\delta} - I(A:B)_{\sigma^\delta}| \leq \varepsilon' \log \text{Tr} P_\delta + 2g(\varepsilon') \]  
\[ \leq \varepsilon' H(\gamma_A(E/\delta)) + 2g(\varepsilon'). \]  
(54)

Since
\[ |I(A:B)_\rho - I(A:B)_{\sigma^\delta}| \leq |I(A:B)_{\rho^\delta} - I(A:B)_{\sigma^\delta}| \]  
\[ + |I(A:B)_\rho - I(A:B)_{\rho^\delta}| + |I(A:B)_{\sigma^\delta} - I(A:B)_{\sigma^\delta}|, \]
continuity bound (49) follows from (53) and (54).

The asymptotic tightness of continuity bound (48) is shown in Remark 11 below.

**Lemma 2.** Let \( P_A \) be a projector in \( \mathfrak{B}(\mathcal{H}_A) \) and \( \omega_{AB} \) be a qc-state (1) with finite \( H(\omega_A) \). Then
\[ -(1 - \tau)H(\bar{\omega}_A) \leq I(A:B)_\omega - I(A:B)_{\omega^\delta} \leq H(\omega_A) - \tau H(\bar{\omega}_A), \]  
(55)

where \( \tau = \text{Tr} P_A \omega_A \) and \( \bar{\omega}_{AB} = \tau^{-1} P_A \otimes I_B \omega_{AB} P_A \otimes I_B \).

**Proof.** The both inequalities in (55) are easily derived from the inequalities
\[ 0 \leq I(A:B)_\omega - \tau I(A:B)_{\omega^\delta} \leq H(\omega_A) - \tau H(\bar{\omega}_A) \]  
(56)

by using nonnegativity of \( I(A:B) \) and upper bound (9).

Note that representation (34) remains valid for an ensemble \( \{ p_i, \rho_i \} \) of any positive trace class operators if we assume that \( H \) and \( I(A:B) \) are

---

Footnote 9: For arbitrary state \( \omega_{AB} \) double inequality (55) holds with additional factors 2 in the left and in the right sides (see Lemma 9 in [29]).
homogeneous extensions of the von Neumann entropy and of the quantum mutual information to the cones of all positive trace class operators and that
\( \chi(\{p_i, \rho_i\}) = H(\bar{\rho}) - \sum_i p_i H(\rho_i) \) provided that \( H(\bar{\rho}) < +\infty \). This shows that the double inequality (56) can be rewritten as follows
\[
0 \leq \chi(\{p_i, \rho_i\}) - \chi(\{p_i, P_A \rho_i P_A\}) \leq H(\bar{\rho}) - H(P_A \bar{\rho} P_A).
\]
The first of these inequalities is easily derived from monotonicity of the quantum relative entropy and concavity of the function \( \eta(x) = -x \log x \).
The second one follows from the definition of the Holevo quantity, since \( H(\rho_i) \geq H(P_A \rho_i P_A) \) for all \( i \) \cite{22}.

By using Proposition 7 and the estimates from \cite{32} one can obtain a continuity bound for the Holevo quantity of ensembles of states of the system composed of \( \ell \) quantum oscillators (described in Section 3.2) under the mean energy constraint. To be consistent with our assumption \( E_0 = 0 \) we will consider shifted Hamiltonian
\[
H_A' = \sum_{i=1}^\ell \hbar \omega_i a_i^+ a_i - \frac{1}{2} \sum_{i=1}^\ell \hbar \omega_i I_A.
\]

By repeating the arguments from the proof of Lemma 18 in \cite{32} with Proposition 7 instead of Meta-Lemmas 16,17 one can obtain the following

**Corollary 6.** Let \( \{p_i, \rho_i\} \) and \( \{q_i, \sigma_i\} \) be countable ensembles of states of the quantum system composed of \( \ell \) oscillators with the average states \( \bar{\rho} \) and \( \bar{\sigma} \) such that \( \text{Tr} H_A' \bar{\rho}, \text{Tr} H_A' \bar{\sigma} \leq E, D_*(\{p_i, \rho_i\}, \{q_i, \sigma_i\}) \leq \varepsilon \leq 1 \). Then
\[
|\chi(\{p_i, \rho_i\}) - \chi(\{q_i, \sigma_i\})| \leq \varepsilon (1 + \frac{\alpha}{1 - \alpha} + 2\alpha) \left[ \sum_{i=1}^\ell \log \left( \frac{E}{\hbar \omega_i} + 1 \right) + \ell \log \frac{e}{\alpha(1-\varepsilon)} \right]
+ \ell \left( 1 + \frac{\alpha}{1 - \alpha} + 2\alpha \right) \tilde{h}_2(\varepsilon) + 2\tilde{h}_2(\alpha\varepsilon) + 2g\left( 1 + \frac{\alpha}{1 - \alpha} \right),
\]
where \( \alpha \in (0, \frac{1}{2}) \), \( \tilde{h}_2(x) = h_2(x) \) for \( x \leq 1/2 \) and \( \tilde{h}_2(x) = 1 \) for \( x \geq 1/2 \), \( g(x) = (x + 1) \log(x + 1) - x \log x \).

Note that the main term in this continuity bound coincides with the main term in the continuity bound for the von Neumann entropy of states of the system of \( \ell \) oscillators with the energy not exceeding \( E \) presented in Lemma 16 in \cite{32}.

**Remark 10.** Lemma 1 in Section 3.2 implies that for large energy \( E \) the main term of the continuity bound in Corollary 6 can be made not greater
than $\epsilon(A_E + o(A_E))$ by appropriate choice of $\alpha$, where

$$A_E = \sum_{i=1}^{\ell} \log \left( \frac{E}{\ell \omega_i} + 1 \right) \approx H(\gamma_A(E)).$$

**Remark 11.** To show the asymptotical tightness of the continuity bound in Proposition 7 for large $E$ it suffices to show this property for the continuity bound in Corollary 6. By Remark 10 this can be done by finding for given $\varepsilon > 0$ and $E > 0$ two ensembles $\{p_i, \rho_i\}$ and $\{q_i, \sigma_i\}$ satisfying the condition of Corollary 6 such that

$$|\chi(\{p_i, \rho_i\}) - \chi(\{q_i, \sigma_i\})| \geq \varepsilon H(\gamma_A(E)).$$

(57)

Let $\{p_i, \rho_i\}$ be any pure state ensemble with the average state $\gamma_A(E)$ and $q_i = p_i, \sigma_i = (1 - \varepsilon)\rho_i + \varepsilon \gamma_A(E)$ for all $i$. Then

$$2D_\varepsilon(\{p_i, \rho_i\}, \{q_i, \sigma_i\}) \leq \sum_{i=1}^{\infty} \|p_i \rho_i - q_i \sigma_i\|_1 = \sum_{i=1}^{\infty} \varepsilon p_i \|\rho_i - \gamma_A(E)\|_1 \leq 2\varepsilon$$

while (57) follows from concavity of the entropy.

## 5 Applications

### 5.1 Tight continuity bounds for the Holevo capacity and for the entanglement-assisted classical capacity of a quantum channel

The **Holevo capacity** of a quantum channel $\Phi : A \to B$ is defined as follows

$$\bar{C}(\Phi) = \sup_{\{p_i, \rho_i\} \in \mathcal{E}(\mathcal{H}_A)} \chi(\{p_i, \Phi(\rho_i)\}),$$

(58)

where the supremum is over all ensembles of input states. This quantity determines the ultimate rate of transmission of classical information through the channel $\Phi$ with non-entangled input encoding, it is closely related to the classical capacity of a quantum channel (see Section 5.2 below) \[13\] \[23\] \[31\].

The **classical entanglement-assisted capacity** of a quantum channel determines an ultimate rate of transmission of classical information when an
entangled state between the input and the output of a channel is used as an additional resource (see details in [13, 23, 31]). By the Bennett-Shor-Smolin-Thaplyal theorem the classical entanglement-assisted capacity of a finite-dimensional quantum channel \( \Phi : A \to B \) is given by the expression

\[
C_{ea}(\Phi) = \sup_{\rho \in \mathcal{S}(\mathcal{H}_A)} I(\Phi, \rho),
\]

in which \( I(\Phi, \rho) \) is the quantum mutual information of the channel \( \Phi \) at a state \( \rho \) defined as follows

\[
I(\Phi, \rho) = I(B:R)_{\Phi \otimes \text{Id}_R(\hat{\rho})},
\]

where \( \mathcal{H}_R \cong \mathcal{H}_A \) and \( \hat{\rho} \) is a pure state in \( \mathcal{S}(\mathcal{H}_{AR}) \) such that \( \hat{\rho}_A = \rho \) [6, 13, 31].

In analysis of variations of the capacities \( \bar{C}(\Phi) \) and \( C_{ea}(\Phi) \) as functions of a channel we will use the operator norm \( \| \cdot \| \) and the diamond norm \( \| \cdot \|_\diamond \) described at the end of Section 2.

Proposition 5A and Corollary 1 imply the following

**Proposition 8.** Let \( \Phi \) and \( \Psi \) be quantum channels from \( A \) to \( B \) and 

\[
g(\varepsilon) = (1 + \varepsilon) h_2(\frac{\varepsilon}{1+\varepsilon}).
\]

Then

\[
|\bar{C}(\Phi) - \bar{C}(\Psi)| \leq \varepsilon \log d_B + 2g(\varepsilon),
\]

where \( \varepsilon = \frac{1}{2} \| \Phi - \Psi \| \) and \( d_B = \dim \mathcal{H}_B \), and

\[
|C_{ea}(\Phi) - C_{ea}(\Psi)| \leq 2\varepsilon \log d + 2g(\varepsilon),
\]

where \( \varepsilon = \frac{1}{2} \| \Phi - \Psi \|_\diamond \) and \( d = \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\} \).

The both continuity bounds (61) and (62) are tight.

**Proof.** For given ensemble \( \{p_i, \rho_i\} \) Proposition 5A shows that

\[
|\chi(\{p_i, \Phi(\rho_i)\}) - \chi(\{p_i, \Psi(\rho_i)\})| \leq \varepsilon_0 \log d_B + 2g(\varepsilon_0),
\]

where \( \varepsilon_0 = \frac{1}{2} \| \Phi - \Psi \|_1 \leq \frac{1}{2} \| \Phi - \Psi \|. \) This and (58) imply (61).

Continuity bounds (62) is derived similarly from Corollary 1 and expression (59), since for any pure state \( \hat{\rho}_{AR} \) in (60) we have

\[
\| \Phi \otimes \text{Id}_R(\hat{\rho}) - \Psi \otimes \text{Id}_R(\hat{\rho}) \|_1 \leq \| \Phi - \Psi \|_\diamond.
\]
To show the tightness of the both continuity bounds assume that $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$, $\Phi$ is the noiseless channel (i.e. $\Phi = \text{Id}_{\mathcal{C}^d}$) and $\Psi$ is the depolarizing channel:

$$\Psi(\rho) = (1 - p)\rho + pd^{-1}I_{\mathcal{C}^d}, \quad p \in [0, 1].$$

Since

$$C(\Psi) = \log d + (1 - pc) \log(1 - pc) + pc \log(p/d),$$
where $c = 1 - 1/d$ [13], $C_{ea}(\Phi) = 2\bar{C}(\Phi) = 2\log d$ and $C_{ea}(\Psi) \leq 2\bar{C}(\Psi)$, we have

$$\bar{C}(\Phi) - \bar{C}(\Psi) = pc \log d + h_2(pc) + pc \log c$$
and

$$C_{ea}(\Phi) - C_{ea}(\Psi) \geq 2pc \log d + 2h_2(pc) + 2pc \log c.$$ 

These relations show the tightness of continuity bound (61) and (62), since it is easy to see that $\|\Phi - \Psi\| \leq \|\Phi - \Psi\|_{\diamond} \leq 2p$. □

5.2 Refinement of the Leung-Smith continuity bounds for classical and quantum capacities of a channel

By the Holevo-Schumacher-Westmoreland theorem the classical capacity of a finite-dimensional channel $\Phi : A \rightarrow B$ is given by the expression

$$C(\Phi) = \lim_{n \rightarrow +\infty} n^{-1}\bar{C}(\Phi^{\otimes n}),$$

where $\bar{C}$ is the Holevo capacity defined in the previous subsection [13 31].

By the Lloyd-Devetak-Shor theorem the quantum capacity of a finite-dimensional channel $\Phi : A \rightarrow B$ is given by the expression

$$Q(\Phi) = \lim_{n \rightarrow +\infty} n^{-1}\bar{Q}(\Phi^{\otimes n}),$$

where $\bar{Q}(\Phi)$ is the maximum of the coherent information $I_c(\Phi, \rho) \triangleq H(\Phi(\rho)) - H(\tilde{\Phi}(\rho))$ over all states $\rho \in \mathcal{S}(\mathcal{H}_A)$ ($\tilde{\Phi}$ is a complementary channel to $\Phi$).

Leung and Smith obtained in [18] the following continuity bounds for classical and quantum capacities of a channel with finite-dimensional output

$$|C(\Phi) - C(\Psi)| \leq 16\varepsilon \log d_B + 4h_2(2\varepsilon),$$

$$|Q(\Phi) - Q(\Psi)| \leq 16\varepsilon \log d_B + 4h_2(2\varepsilon),$$

29
where \( \varepsilon = \frac{1}{2} \| \Phi - \Psi \|_\diamond \) and \( d_B = \text{dim} \mathcal{H}_B \). By using Winter’s tight continuity bound \( (6) \) for the conditional entropy (instead of the original Alicki-Fannes continuity bound) in the Leung-Smith proof one can replace the main terms in \( (65) \) and \( (66) \) by \( 4\varepsilon \log d_B \). By using Proposition 3A one can replace the main terms in \( (65) \) and \( (66) \) by \( 2\varepsilon \log d_B \) (which gives tight continuity bound for the quantum capacity and close-to-tight continuity bound for the classical capacity).

**Proposition 9.** Let \( \Phi \) and \( \Psi \) be channels from \( A \) to \( B \). Then

\[
|C(\Phi) - C(\Psi)| \leq 2\varepsilon \log d_B + 2g(\varepsilon),
\]

where \( \varepsilon = \frac{1}{2} \| \Phi - \Psi \|_\diamond \), \( d_B = \text{dim} \mathcal{H}_B \) and \( g(\varepsilon) = (1 + \varepsilon)h_2 \left( \frac{\varepsilon}{1+\varepsilon} \right) \).

Continuity bound \( (68) \) is tight, continuity bound \( (67) \) is close-to-tight (up to the factor 2 in the main term).

**Proof.** Since

\[
\bar{C}(\Phi^\otimes n) = \sup \chi(\{\pi_i, \Phi^\otimes n(\rho_i)\}),
\]

where the supremum is over all ensembles \( \{\pi_i, \rho_i\} \) of states in \( \mathcal{S}(\mathcal{H}_A^\otimes n) \), continuity bound \( (67) \) is obtained by using Lemma 12 in [18], representation \( (34) \) and Proposition 3A in Section 3.3.

To prove continuity bound \( (68) \) note that the coherent information can be represented as follows

\[
I_c(\Phi, \rho) = I(B:R)_{\Phi \otimes \text{Id}_R}(\hat{\rho}) - H(\rho),
\]

where \( \hat{\rho} \in \mathcal{S}(\mathcal{H}_A R) \) is a purification a state \( \rho \). Hence for arbitrary quantum channels \( \Phi \) and \( \Psi \), arbitrary \( n \) and any state \( \rho \) in \( \mathcal{S}(\mathcal{H}_A^\otimes n) \) we have

\[
I_c(\Phi^\otimes n, \rho) - I_c(\Psi^\otimes n, \rho) = I(B^n:R^n)_{\Phi^\otimes n \otimes \text{Id}_R^n}(\hat{\rho}) - I(B^n:R^n)_{\Psi^\otimes n \otimes \text{Id}_R^n}(\hat{\rho})
\]

where \( \hat{\rho} \in \mathcal{S}(\mathcal{H}_A^\otimes n) \) is a purification of the state \( \rho \). This representation, Proposition 3A in Section 3.3 and Lemma 12 in [18] imply the continuity bound for the quantum capacity.

The tightness of the continuity bound for the quantum capacity can be shown by using the erasure channels \( (33) \) from \( d \)-dimensional system \( A \) to \( (d + 1) \)-dimensional system \( B \). It is known that \( Q(\Phi_\rho) = (1 - 2p) \log d \) for

\[10\text{It is assumed that expressions } (63) \text{ and } (64) \text{ remain valid in the case } \text{dim} \mathcal{H}_A = +\infty\text{.} \]
$p \leq 1/2$ and $Q(\Phi_p) = 0$ for $p \geq 1/2$ [31]. Hence $Q(\Phi_0) - Q(\Phi_p) = 2p \log d$ for $p \leq 1/2$. By noting that $\|\Phi_0 - \Phi_p\|_{\infty} \leq 2p$ we see that continuity bound (68) is tight (for large $d$).

The proof of tightness of continuity bound (61) for the Holevo capacity shows that the main term in (67) is close to the optimal one up to the factor $2$, since $C(\Phi)$ coincides with $\bar{C}(\Phi)$ for depolarizing channel $\Phi$. □

5.3 Continuity bounds for classical capacities of infinite-dimensional channels with energy constraints

When we consider transmission of classical information over infinite dimensional quantum channels we have to impose the energy constraint on states used for coding information. For a single channel $\Phi : A \to B$ the energy constraint is determined by the linear inequality

$$\text{Tr} H_A \rho \leq E, \quad E > 0,$$

(69)

where $H_A$ is the Hamiltonian of the input system $A$. For $n$ copies of this channel the energy constraint is given by the inequality

$$\text{Tr} \rho^{(n)} H_{A^n} \leq nE,$$

(70)

where $\rho^{(n)}$ is a state of the system $A^n$ ($n$ copies of $A$) and

$$H_{A^n} = H_A \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes H_A$$

(71)

is the Hamiltonian of the system $A^n$.

An operational definition of the classical capacity of a quantum channel with linear constraint can be found in [14]. If only nonentangled input encoding is used then the ultimate rate of transmission of classical information through the channel $\Phi$ with the constraint (70) on mean energy of a code is determined by the Holevo capacity

$$\bar{C}(\Phi, H_A, E) = \sup_{\text{Tr} H_A \bar{\rho} \leq E} \chi(\{p_i, \Phi(\rho_i)\}), \quad \bar{\rho} = \sum_i p_i \rho_i$$

(72)

(the supremum is over all input ensembles $\{p_i, \rho_i\}$ such that $\text{Tr} H_A \bar{\rho} \leq E$). By the Holevo-Schumacher-Westmoreland theorem adapted to constrained
channels ([14], Proposition 3]), the classical capacity of the channel $\Phi$ with constraint (70) is given by the following regularized expression

$$C(\Phi, H_A, E) = \lim_{n \to +\infty} n^{-1} C(\Phi^\otimes n, H_{A^n}, nE),$$

where $H_{A^n}$ is defined in (71). If $\bar{C}(\Phi^\otimes n, H_{A^n}, nE) = n\bar{C}(\Phi, H_A, E)$ for all $n$ then

$$C(\Phi, H_A, E) = \bar{C}(\Phi, H_A, E)$$

(73)
i.e. the classical capacity of the channel $\Phi$ coincides with its Holevo capacity.

Note that (73) holds for many infinite dimensional channels [13]. Recently it is shown that (73) holds if $\Phi$ is a gauge covariant or contravariant Gaussian channel and $H_A = \sum_{ij} \epsilon_{ij} a_i^\dagger a_j$ - gauge covariant Hamiltonian (here $[\epsilon_{ij}] -$ is a positive matrix) [10, 11].

The following proposition presents estimates for differences between the Holevo capacities and between the classical capacities of channels $\Phi$ and $\Psi$ with finite energy amplification factors which means that

$$\sup_{\text{Tr} H_A \rho \leq E} \text{Tr} H_B (\Phi(\bar{\rho})) \leq kE$$ and $$\sup_{\text{Tr} H_A \rho \leq E} \text{Tr} H_B (\Psi(\bar{\rho})) \leq kE$$

(74)

for some finite $k$. Note that any channels produced in a physical experiment satisfy condition (74).

**Proposition 10.** Let $\Phi$ and $\Psi$ be channels from $A$ to $B$ satisfying condition (74), $\varepsilon = \frac{1}{2} || \Phi - \Psi ||_\diamond$. If the Hamiltonian $H_B$ of system $B$ satisfies conditions (25) and (26), $\varepsilon' \in (\varepsilon, 1]$ and $\delta = \frac{\varepsilon'}{1+\varepsilon}$ then

$$|\bar{C}(\Phi, H_A, E) - \bar{C}(\Psi, H_A, E)| \leq (\varepsilon' + 2\delta)H(\gamma_B(kE/\delta)) + 2g(\varepsilon') + 2h_2(\delta),$$

(75)

and

$$|C(\Phi, H_A, E) - C(\Psi, H_A, E)| \leq (2\varepsilon' + 4\delta)H(\gamma_B(kE/\delta)) + 2g(\varepsilon') + 4h_2(\delta),$$

(76)

where $\gamma_B(E)$ is the Gibbs state in system $B$.

Continuity bound (75) is asymptotically tight for large $E$, continuity bound (76) is close-to-tight (up to the factor 2 in the main term).

**Proof.** Since condition (74) implies $\text{Tr} H_B \Phi(\bar{\rho}) \leq kE$ and $\text{Tr} H_B \Psi(\bar{\rho}) \leq kE$ for any ensemble $\{p_i, \rho_i\}$ such that $\text{Tr} H_A \bar{\rho} \leq E$, continuity bound (73) is

\[11\]The gauge covariance condition for $H_A$ can be replaced by the condition (18) in [11].

32
obtained by using Winter’s type continuity bound for the Holevo quantity (Proposition 7).

To prove continuity bound (76) note that
\[
\bar{C}(\Phi^\otimes n, H_{A^n}, nE) = \sup \chi(\{p_i, \Phi^\otimes n(\rho_i)\}),
\]
where the supremum is over all ensembles \(\{p_i, \rho_i\}\) of states in \(\mathcal{S}(\mathcal{H}_{A^n})\) with the average state \(\bar{\rho}\) satisfying the condition
\[
\operatorname{Tr} H_{A^n} \bar{\rho} = \sum_{k=1}^n \operatorname{Tr} H_{A^k} \bar{\rho}_{A_k} \leq nE, \quad \bar{\rho}_{A_k} = \operatorname{Tr}_{A^{n-k}} \bar{\rho}.
\]
(77)

Since condition (74) implies
\[
\sum_{k=1}^n \operatorname{Tr} H_B \Phi(\bar{\rho}_{A_k}) \leq nkE \quad \text{and} \quad \sum_{k=1}^n \operatorname{Tr} H_B \Psi(\bar{\rho}_{A_k}) \leq nkE
\]
for any ensemble \(\{p_i, \rho_i\}\) satisfying condition (77), continuity bound (76) is obtained by using representation (34), Proposition 3B and the corresponding analog of Lemma 12 in [18].

The tightness of the continuity bound (75) can be shown by using the erasure channels (33) from the system \(A\) composed of \(\ell\) quantum oscillators to any its one-dimensional extension \(B\). These channels satisfy condition (74) with \(k = 1\). It is easy to see that \(\bar{C}(\Phi_p, H_A, E) = (1 - p)H(\gamma(E))\), where \(\gamma(E)\) is the Gibbs state corresponding to the energy \(E\). Hence
\[
|\bar{C}(\Phi_0, H_A, E) - \bar{C}(\Phi_p, H_A, E)| = p\gamma(H(\gamma(E)))
\]
(78)

By Remark 10 in this case the main term of (75) can be made not greater than \(\varepsilon[H(\gamma(E)) + o(H(\gamma(E)))]\) for large \(E\) by appropriate choice of \(\varepsilon'\). So, the asymptotic tightness of continuity bound (75) follows from (78), since \(\|\Phi_0 - \Phi_p\|_\diamond \leq 2p\).

The above example also shows that the main term in continuity bound (76) is close to the optimal one up to the factor 2, since \(C(\Phi_p, H_A, E)\) coincides with \(\bar{C}(\Phi_p, H_A, E)\) for any \(p\). □

An operational definition of the entanglement-assisted classical capacity of an infinite dimensional quantum channel with energy constraint (69) is given in [14]. By the Bennett-Shor-Smolin-Thaplyal theorem adapted to
constrained channels \([14, \text{Proposition 4}]\) the entanglement-assisted classical capacity an infinite dimensional channel \(\Phi\) with the energy constraint \((69)\) determined by a Hamiltonian \(H_A\) satisfying condition \((25)\) is given by the expression

\[
C_{ea}(\Phi, H_A, E) = \sup_{\text{Tr}H_A\rho \leq E} I(\Phi, \rho),
\]

where \(I(\Phi, \rho)\) is the quantum mutual information of the channel \(\Phi\) at a state \(\rho\) defined by \((60)\).

Proposition 2 implies the following

**Proposition 11.** Let \(\Phi\) and \(\Psi\) be channels from \(A\) to \(B\), \(\varepsilon = \frac{1}{2}\|\Phi - \Psi\|_\diamond\), \(\varepsilon' \in (\varepsilon, 1]\), \(\delta = \frac{\varepsilon' - \varepsilon}{1 + \varepsilon}\) and \(\gamma_X(E)\) is the Gibbs state in system \(X = A, B\).

A) If the Hamiltonian \(H_A\) satisfies conditions \((25)\) and \((26)\) then

\[
|C_{ea}(\Phi, H_A, E) - C_{ea}(\Psi, H_A, E)| \leq (2\varepsilon' + 4\delta)H(\gamma_A(E/\delta)) + 2g(\varepsilon') + 4h_2(\delta). \tag{79}
\]

B) If the channels \(\Phi\) and \(\Psi\) satisfies condition \((74)\) and the Hamiltonian \(H_B\) satisfies conditions \((25)\) and \((26)\) then

\[
|C_{ea}(\Phi, H_A, E) - C_{ea}(\Psi, H_A, E)| \leq (2\varepsilon' + 4\delta)H(\gamma_B(kE/\delta)) + 2g(\varepsilon') + 4h_2(\delta). \tag{80}
\]

**Continuity bounds** \((79)\) and \((80)\) are asymptotically tight for large \(E\).

Note that continuity bound \((79)\) holds for arbitrary channels \(\Phi\) and \(\Psi\).

**Proof.** A) Let \(\mathcal{H}_R \cong \mathcal{H}_A\) and \(H_R\) be an operator in \(\mathcal{H}_R\) unitarily equivalent to \(H_A\). For any state \(\rho\) satisfying the condition \(\text{Tr}H_A\rho \leq E\) there exists a purification \(\hat{\rho} \in \mathcal{S}(\mathcal{H}_{AR})\) such that \(\text{Tr}H_R\hat{\rho} \leq E\). Since

\[
I(\Phi, \rho) = I(B:R)_{\sigma} \quad \text{and} \quad I(\Psi, \rho) = I(B:R)_{\varsigma},
\]

where \(\sigma = \Phi \otimes \text{Id}_R(\hat{\rho})\) and \(\varsigma = \Psi \otimes \text{Id}_R(\hat{\rho})\) are states in \(\mathcal{S}(\mathcal{H}_{BR})\) such that \(\text{Tr}H_R\sigma_{SR}, \text{Tr}H_{R\Sigma R} \leq E\) and \(\|\sigma - \varsigma\|_1 \leq \|\Phi - \Psi\|_\diamond\), Proposition 2 shows that \(|I(\Phi, \rho) - I(\Psi, \rho)|\) is upper bounded by the right hand side of \((79)\).

B) Continuity bound \((80)\) is obtained similarly from Proposition 2, since in this case \(\text{Tr}H_{RSB}, \text{Tr}H_{RSB} \leq kE\).

The tightness of the both continuity bounds is also shown by using the erasure channels \((33)\) from the system \(A\) composed of \(\ell\) quantum oscillators to any its one-dimensional extension \(B\). It suffices only to note that \(C_{ea}(\Phi_0, H_A, E) = 2H(\gamma(E))\) and \(C_{ea}(\Phi_p, H_A, E) \leq 2(1 - p)\tilde{H}(\gamma(E))\) and to repeat the arguments from the proof of Proposition 10. \(\Box\)
Since condition (25) implies \( \lim_{\delta \to +0} \delta H(\gamma(E/\delta)) = 0 \) [27, Pr.1], we obtain from Propositions 10 and 11 the following observations.

**Corollary 7.** Let \( \mathfrak{F}(A,B) \) be the set of all quantum channels from \( A \) to \( B \) equipped with the diamond norm topology.

A) If the Hamiltonian \( H_B \) of system \( B \) satisfies conditions (25) then the functions \( \Phi \mapsto \bar{C}(\Phi, H_A, E) \), \( \Phi \mapsto C(\Phi, H_A, E) \) and \( \Phi \mapsto C_{ca}(\Phi, H_A, E) \) are uniformly continuous on any subset of \( \mathfrak{F}(A,B) \) consisting of channels with bounded energy amplification factor.

B) If the Hamiltonian \( H_A \) of system \( A \) satisfies conditions (25) then the function \( \Phi \mapsto C_{ca}(\Phi, H_A, E) \) is uniformly continuous on \( \mathfrak{F}(A,B) \).

A drawback of Corollary 7 is the use of the diamond norm topology on the set of infinite-dimensional channels, since this topology is too strong for analysis of real variations of such channels.\footnote{There are channels with close physical parameters having large diamond norm of the difference [33]. This is explained, briefly speaking, by the fact that the diamond norm topology on the set of channels corresponds to the uniform operator topology on the set of Stinespring isometries [16], see the remark in [28, Section 8.2].} More preferable topology on the set of infinite-dimensional quantum channels is the strong convergence topology defined by the family of seminorms \( \Phi \mapsto \|\Phi(\rho)\|_1, \rho \in \mathcal{S}(H_A) \). Some assertions of Corollary 7 are generalized to the case of this topology, e.g., Proposition 11 in [28] asserts global continuity of the function \( \Phi \mapsto C_{ca}(\Phi, H_A, E) \) with respect to the strong convergence topology if the Hamiltonian \( H_A \) satisfies conditions (25). The most difficult open problem is to prove the strong convergence topology version of Corollary 7A for the classical capacity (because of the regularization in its definition). Another interesting task is to prove the analogue of Corollary 7B for the capacities \( \bar{C}(\Phi, H_A, E) \) and \( C(\Phi, H_A, E) \).

**Appendix: the proof of Proposition 4**

A) To show that \( D_s(\mu,\nu) \leq D_{ehs}(\mu,\nu) \) for any ensembles \( \mu = \{p_i, \rho_i\} \) and \( \nu = \{q_i, \sigma_i\} \) it suffices to note that

\[
\sum_{i,j} \|P_{ij}\rho_i - Q_{ij}\sigma_j\|_1 = 2D_0(\mu',\nu'),
\]

where \( \mu' = \{P_{ij}, \rho_i\}_{ij} \in E(\mu) \) and \( \nu' = \{Q_{ij}, \sigma_j\}_{ij} \in E(\nu) \).
Since \( D_{\text{ehs}}(\mu, \nu) \leq D_0(\mu, \nu) \), the inequality \( D_{\text{ehs}}(\mu, \nu) \leq D_*(\mu, \nu) \) can be proved by showing that the metric \( D_{\text{ehs}} \) does not change under permutations of states and under splitting of states of the both ensembles.

The invariance of \( D_{\text{ehs}} \) under permutations follows from definition (37): permutations of states of the ensemble \( \{p_i, \rho_i\} \) correspond to permutations of rows of the matrices \( P_{ij} \) and \( Q_{ij} \), permutations of states of the ensemble \( \{q_i, \sigma_i\} \) correspond to permutations of columns of these matrices. So, by symmetry, it suffices to show that

\[
D_{\text{ehs}}(\mu, \nu) = D_{\text{ehs}}(\mu', \nu) \quad (81)
\]

for any ensembles \( \mu = \{p_i, \rho_i\}, \nu = \{q_i, \sigma_i\} \) and the ensemble \( \mu' = \{p'_i, \rho'_i\} \) obtained by splitting of the first state of \( \mu \) in which \( \rho'_1 = \rho'_2 = \rho_1, \ p'_1 = kp_1 \), \( p'_2 = (1-k)p_1 \ (k \in [0, 1]) \) and \( \rho'_i = \rho_{i-1}, \ p'_i = p_{i-1} \) for \( i > 2 \).

For given \( \varepsilon > 0 \) let \( P_{ij} \) and \( Q_{ij} \) be joint probability distributions such that

\[
\sum_{i,j} \| P_{ij} \rho_i - Q_{ij} \sigma_j \|_1 \leq 2D_{\text{ehs}}(\mu, \nu) - \varepsilon, \quad \sum_j P_{ij} = p_i, \quad \sum_i Q_{ij} = q_j. \quad (82)
\]

Let \( P'_{ij} \) be the matrix obtained from the matrix \( P_{ij} \) by replacing its first row \( [P_{11}, P_{12}, ...] \) by the block

\[
\begin{bmatrix}
  kP_{11}, kP_{12}, ... \\
  \bar{k}P_{11}, \bar{k}P_{12}, ...
\end{bmatrix}, \quad \bar{k} = k - 1,
\]

and \( Q'_{ij} \) the matrix obtained from the matrix \( Q_{ij} \) by the similar way. Then

\[
2D_{\text{ehs}}(\mu', \nu) \leq \sum_{i,j} \| P'_{ij} \rho'_i - Q'_{ij} \sigma'_j \|_1 = \sum_j \| kP_{ij} \rho_i - kQ_{ij} \sigma_j \|_1 + \sum_j \| \bar{k}P_{ij} \rho_i - \bar{k}Q_{ij} \sigma_j \|_1 + \sum_{i>1,j} \| P_{ij} \rho_i - Q_{ij} \sigma_j \|_1
\]

\[
= \sum_{i,j} \| P_{ij} \rho_i - Q_{ij} \sigma_j \|_1 \leq 2D_{\text{ehs}}(\mu, \nu) - \varepsilon,
\]

which proves “\( \geq \)” in (81).

For given \( \varepsilon > 0 \) let \( P'_{ij} \) and \( Q'_{ij} \) be joint probability distributions for which the relation similar to (82) holds for the ensembles \( \mu' \) and \( \nu \). Let \( P_{ij} \) be the matrix obtained from the matrix \( P'_{ij} \) by replacing its first two rows
by the row \([P_i' + P_j', P_1', P_2', ..., P_n']\) and \(Q_{ij}\) the matrix obtained from the matrix \(Q'_{ij}\) by the similar way. Then

\[
2D_{\text{ehs}}(\mu, \nu) \leq \sum_{i,j} \|P_{ij}\rho_i - Q_{ij}\sigma_j\|_1 = \sum_{i,j} \|(P_{1j}' + P_{2j}')\rho_1 - (Q_{1j}' + Q_{2j}')\sigma_j\|_1
\]

\[
+ \sum_{i>2,j} \|P_{ij}'\rho_i - Q_{ij}'\sigma_j\|_1 \leq \sum_{i,j} \|P_{ij}'\rho_i - Q_{ij}'\sigma_j\|_1 \leq 2D_{\text{ehs}}(\mu', \nu) - \varepsilon,
\]

which shows that "\(\leq\)" holds in (81).

B) It is shown in [24] that convergence of a sequence \(\{p_i^n, \rho_i^n\}\) to an ensemble \(\{p_i^0, \rho_i^0\}\) with respect to the metric \(D_{\text{ehs}}\) implies

\[
\lim_{n \to \infty} \sum_i p_i^n f(\rho_i^n) = \sum_i p_i^0 f(\rho_i^0)
\]

for any uniformly continuous bounded function \(f\) on \(\mathcal{G}(\mathcal{H})\). By Theorem 6.1 in [25] this means that the \(D_{\text{ehs}}\)-convergence is not weaker than the weak convergence. So, by assertion A it suffices to show that the \(D_{\varepsilon}\)-convergence is not stronger than the weak convergence.

Let \(\{\mu_n = \{p_i^n, \rho_i^n\}\}\) be a sequence weakly converging to an ensemble \(\mu_0 = \{q_i, \sigma_i\}\) and \(\varepsilon > 0\) be arbitrary. By adding any states with zero probabilities we may assume that all the ensembles \(\mu_n\) and \(\mu_0\) are countable. Let \(m\) be such that \(\sum_{i>m} q_i < \varepsilon\) and \(U_1, ..., U_m\) mutually disjoint ball vicinities of the states \(\sigma_1, ..., \sigma_m\) having radii \(\leq \varepsilon\) such that \(\sum_{i: \sigma_i \in U_k} q_i < q_k + \varepsilon/m\) and the boundary of \(U_k\) does not contain states of \(\mu_0\) for all \(k = 1, 2, ..., m\).

By the weak convergence of the sequence \(\{\mu_n\}\) to the ensemble \(\mu_0\) there is \(n_\varepsilon\) such that \(\left| \sum_{i: q_i^n \in U_k} p_i^n - \sum_{i: \sigma_i \in U_k} q_i \right| < \varepsilon/m\) and hence

\[
\left| \sum_{i: q_i^n \in U_k} p_i^n - q_k \right| \leq \sum_{i: q_i^n \in U_k} p_i^n - \sum_{i: \sigma_i \in U_k} q_i + \left| \sum_{i: \sigma_i \in U_k} q_i - q_k \right| < 2\varepsilon/m
\]

for all \(n \geq n_\varepsilon\) and all \(k = 1, 2, ..., m\). So, for any \(n \geq n_\varepsilon\) in each set \(U_k\) one can take \(l_k < +\infty\) states from the ensemble \(\mu_n\) whose total probability is \(\varepsilon'-\)close to \(q_k\), where \(\varepsilon' = 3\varepsilon/m\). Denote these states and the corresponding probabilities respectively by \(q_{k_1}, ..., q_{k_l}\) and \(r_{k_1}^{l_1}, ..., r_{k_l}^{l_l}\). Let \(t_k = \sum_{i=1}^{l_k} r_i^k\), so that \(\left| t_k - q_k \right| < 3\varepsilon/m\). The states of the ensemble \(\mu_n\) not included in the above collections (taken in any order) and their probabilities denote respectively by \(q_{k_0}, ..., q_{l_0}^0\) and \(r_{k_0}^{l_0}, ..., r_{l_0}^{l_0}\), where \(l_0 = +\infty\).
Let $s_i^k = r_i^k q_k t_i^k$, $c_i^k = \sigma_k$ for $k = 1, 2, \ldots, m_i$, $i = 1, 2, \ldots, l_k$ and $c_i^0 = \sigma_{m+i}$, $s_i^0 = q_{m+i}$ for $i = 1, 2, \ldots, l_0$. Let $\mu'_n = \{r_i^k, g_i^k\}_{ki}$ and $\mu'_0 = \{s_i^k, c_i^k\}_{ki}$, where $k$ runs from 0 to $m$ and $i$ runs from 1 to $l_k$. Since $\mu'_n \in E(\mu_n)$ and $\mu'_0 \in E(\mu_0)$, we have

$$2D_*(\mu_n, \mu_0) \leq 2D_0(\mu'_n, \mu'_0) = \sum_{k=1}^m \sum_{i=1}^{l_k} \|r_i^k g_i^k - s_i^k c_i^k\|_1 + \sum_{i=1}^{l_0} \|r_i^0 g_i^0 - s_i^0 c_i^0\|_1$$

$$\leq \sum_{k=1}^m \sum_{i=1}^{l_k} s_i^k \|g_i^k - \sigma_k\|_1 + \sum_{k=1}^m \sum_{i=1}^{l_k} |s_i^k - r_i^k| + \sum_{i=1}^{l_0} (r_i^0 + s_i^0).$$

The first sum in the right hand side is less than $\varepsilon$, since $\|g_i^k - \sigma_k\|_1 < \varepsilon$ for all $k = 1, 2, \ldots, m_i$, $i = 1, 2, \ldots, l_k$. The second sum is upper bounded by $\sum_{k>0} q_k + 1 - \sum_{k=1}^m t_k \leq \varepsilon + 1 - \sum_{k=1}^m q_k + \sum_{k=1}^m |q_k - t_k| < \varepsilon + \varepsilon + 3\varepsilon = 5\varepsilon$.

Hence $2D_*(\mu_n, \mu_0) \leq 9\varepsilon$. This shows that the sequence $\{\mu_n\}$ converges to the ensemble $\mu_0$ in the metric $D_*$. □

I am grateful to A.Winter for valuable communication and for several technical tricks essentially used in this work. I am also grateful to A.S.Holevo and G.G.Amosov for useful discussion.

The research is funded by the grant of Russian Science Foundation (project No 14-21-00162).

References

[1] R.Alicki, M.Fannes, ”Continuity of quantum conditional information”, Journal of Physics A: Mathematical and General, V.37, N.5, L55-L57 (2004); arXiv: quant-ph/0312081.

[2] K.M.R.Audenaert, ”A sharp continuity estimate for the von Neumann entropy”, J. Math. Phys. A: Math. Theor. 40(28), 8127-8136 (2007).

[3] K.M.R. Audenaert, J. Eisert, ”Continuity bounds on the quantum relative entropy”, J. Math. Phys. 46, 102104 (2005); quant-ph/0503218.
[4] K.M.R. Audenaert, J. Eisert, "Continuity bounds on the quantum relative entropy-II", J. Math. Phys. 52, 112201 (2011); arXiv:1105.2656.

[5] K.M.R. Audenaert "Quantum Skew Divergence", J. Math. Phys. 55, 112202 (2014); arXiv:1304.5935.

[6] C.H.Bennett, P.W.Shor, J.A.Smolin, A.V.Thapliyal "Entanglement-assisted classical capacity of noisy quantum channel", Phys. Rev. Lett. V.83. 3081-3084 (1999).

[7] P.Billingsley, "Convergence of probability measures", John Willey and Sons. Inc., New York-London-Sydney-Toronto, 1968.

[8] G.F.Dell'Antonio "On the limits of sequences of normal states", Commun. Pure Appl. Math. V.20, 413-430, (1967).

[9] M. Fannes, "A continuity property of the entropy density for spin lattice systems", Commun. Math. Phys. V.31, 291-294 (1973).

[10] V.Giovannetti, A.S.Holevo, R.Garcia-Patron, "A solution of Gaussian optimizer conjecture for quantum channels", Commun. Math. Phys., V.334, N.3, 1553-1571 (2015); arXiv:1312.2251.

[11] A.S.Holevo, "On the constrained classical capacity of infinite-dimensional covariant quantum channels", J. Math. Phys., V.57, N.1, 015203, 11 pp (2016); arXiv:1409.8085.

[12] A.S.Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel", Probl. Inf. Transm. (USSR) V.9, 177-183 (1973).

[13] A.S.Holevo "Quantum systems, channels, information. A mathematical introduction", Berlin, DeGruyter, 2012.

[14] A.S.Holevo "Classical capacities of quantum channels with constrained inputs", Probability Theory and Applications. V.48. N.2. 359-374 (2003).

[15] A.S.Holevo, M.E.Shirokov "Continuous ensembles and the $\chi$-capacity of infinite dimensional channels", Theory of Probability and its Applications, V.50, N.1, P.8698 (2005); arXiv:quant-ph/0408176.
[16] D. Kretschmann, D. Schlingemann, R. F. Werner, "A Continuity Theorem for Stinespring’s Dilation”, arXiv:0710.2495.

[17] A. A. Kuznetsova, "Quantum conditional entropy for infinite-dimensional systems”, Theory of Probability and its Applications, V. 55, N. 4, 709-717 (2011).

[18] D. Leung, G. Smith, "Continuity of quantum channel capacities”, Commun. Math. Phys., V. 292, 201-215 (2009).

[19] N. Li, S. Luo, "Classical and quantum correlative capacities of quantum systems”, Phys. Rev. A 84, 042124 (2011).

[20] E. H. Lieb, M. B. Ruskai, "Proof of the strong suadditivity of quantum mechanical entropy”, J. Math. Phys. V. 14. 1938 (1973).

[21] G. Lindblad "Entropy, information and quantum measurements”, Comm. Math. Phys. V. 33. 305-322 (1973).

[22] G. Lindblad "Expectation and Entropy Inequalities for Finite Quantum Systems”, Comm. Math. Phys. V. 39. N. 2. 111-119 (1974).

[23] M. A. Nielsen, I. L. Chuang "Quantum Computation and Quantum Information”, Cambridge University Press, 2000.

[24] O. Oreshkov, J. Calsamiglia, "Distinguishability measures between ensembles of quantum states”, Phys. Rev. A 79, 032336 (2009); arXiv:0812.3238.

[25] K. Parthasarathy "Probability measures on metric spaces”, Academic Press, New York and London, 1967;

[26] D. Reeb, M. M. Wolf, "Tight bound on relative entropy by entropy difference”, IEEE Trans. Inf. Theory 61, 1458-1473 (2015); arXiv:1304.0036.

[27] M. E. Shirokov, "Entropic characteristics of subsets of states I”, Izvestiya: Mathematics, V. 70, N. 6, 1265-1292 (2006); arXiv:quant-ph/0510073.

[28] M. E. Shirokov, "Measures of correlations in infinite-dimensional quantum systems”, arXiv:1506.06377, to appear in Sbornik: Mathematics.
[29] M.E. Shirokov, "Squashed entanglement in infinite dimensions", J. Math. Phys. V.57, N.3, 032203, 22 pp. (2016); arXiv:1507.08964.

[30] A. Wehrl, "General properties of entropy", Rev. Mod. Phys. 50, 221-250, (1978).

[31] M.M. Wilde, "From Classical to Quantum Shannon Theory", arXiv:1106.1445 (v.6).

[32] A. Winter, "Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints", arXiv:1507.07775 (v.6), to appear in Commun.Math.Phys.

[33] A. Winter, private communication.