Uniform continuity bounds for information characteristics of quantum channels depending on input dimension and on input energy

M.E. Shirokov

Steklov Mathematical Institute, Moscow Institute of Physics and Technology, email:msh@mi.ras.ru

October 2018

Abstract. We obtain continuity bounds for basic information characteristics of quantum channels depending on their input dimension (if it is finite) and on the input energy bound (if the input dimension is infinite). We pay a special attention to the case of a multi-mode quantum oscillator as an input system.

First, we prove continuity bounds for the output conditional mutual information for a single channel and for n copies of a channel.

Then we obtain estimates for variation of the output Holevo quantity with respect to simultaneous variations of a channel and of an input ensemble.

As a result, tight and close-to-tight continuity bounds for basic capacities of quantum channels depending on the input dimension are obtained. They complement the Leung-Smith continuity bounds depending on the output dimension.

Finally, we obtain tight and close-to-tight continuity bounds for basic capacities of infinite-dimensional energy-constrained channels with respect to the energy-constrained Bures distance generating the strong convergence of quantum channels.

1. Introduction

In study of informational properties of quantum channels it is necessary to take into account all possible perturbations of quantum states and quantum channels unavoidable in real physical scenarios. Mathematically, this means that we have to quantitatively analyse continuity of basic information characteristics (in particular, channel capacities) with respect to appropriate metrics on the sets of quantum states and quantum channels.

Leung and Smith obtained in [1] (uniform) continuity bounds (estimates for variations) for basic capacities of quantum channels depending of their output

dimension. The appearance of the output dimension in these and some other continuity bounds for information characteristics of a quantum channel is natural, since such characteristics are typically expressed via entropic quantities of output states of a channel (so, application of Fannes' type continuity bounds gives the factor proportional to the logarithm of the output dimension [2, 3, 4, 5]).

At the same time, it is the input dimension of a channel that determines the range of its information capacities, while the formal output dimension may be substantially greater than the real dimension of a channel output. So, it is reasonable to conjecture that the input dimension also determines the rate of variations of capacities regardless of the output dimension.

Speaking about capacities of channels between infinite-dimensional quantum systems we have to impose energy constraints on states used for encoding information [6, 7, 8]. In [9, 10] continuity bounds for classical and quantum capacities are obtained for energy-constrained infinite-dimensional channels with *bounded energy amplification factor*. In fact, these continuity bounds depend on the output average energy of states used for encoding information. Nevertheless, the range of information capacities is completely determined by the input average energy of code-states, which plays, roughly speaking, the role of input dimension [5]. So, we may expect that the input energy bound also determines the rate of variations of capacities and of other characteristics of infinite-dimensional channels regardless of the output energy (which may be unbounded or equal to $+\infty$).

In this paper we confirm both these conjectures by deriving continuity bounds for several information characteristics of quantum channels (in particular, for all basic capacities) depending on their input dimension (if it is finite) and on the input energy bound (if the input dimension is infinite) provided that the Hamiltonian H_A of the input system satisfies the condition

$$\lim_{\lambda \to 0^+} \left[\mathrm{Tr}\, e^{-\lambda H_A} \right]^{\lambda} = 1,$$

which holds, in particular, if the input system is a multi-mode quantum oscillator [6, 11]. This case plays a central role in quantum information theory. It is considered separately after formulations of general results.

We will consider the case of finite input dimension and the case of finite input energy simultaneously excepting the last Sections 7 and 8 devoted, respectively, to the first and to the second case.

In Section 3 we describe appropriate metrics on the set of quantum channels. In particular, we introduce the *energy-constrained Bures distance* between quantum channels used in the "finite input energy" part of the paper.

Then we prove special continuity bounds for the extented quantum conditional mutual information (Lemma 4 in Section 4), which are our main technical tools.

In Section 5 we obtain continuity bounds for the output conditional mutual information $I(B : D|C)_{\Phi \otimes \mathrm{Id}_{CD}(\rho)}$ with respect to simultaneous variations of a channel

$\Phi : A \to B$ and of an input state ρ_{ACD}. We also derive continuity bound for the function $\Phi \mapsto I(B^n : D|C)_{\Phi^{\otimes n} \otimes \mathrm{Id}_{CD}(\rho)}$ for any natural n by using the Leung-Smith telescopic trick.

In Section 6 we analyse continuity properties of the output Holevo quantity $\chi(\Phi(\mu))$ – the Holevo quantity of the ensemble $\Phi(\mu)$ obtained by action of a channel Φ on a given (discrete or continuous) ensemble μ of input states. We obtain estimates for variation of $\chi(\Phi(\mu))$ with respect to simultaneous variations of a channel Φ and of an input ensemble μ.

In Section 7 the above results are applied to obtain tight and close-to-tight continuity bounds for basic capacities of quantum channels depending on their input dimension. They complement the Leung-Smith continuity bounds (depending on the output dimension).

In Section 8 we obtain tight and close-to-tight continuity bounds for basic capacities of infinite-dimensional energy-constrained channels with respect to the energy-constrained Bures distance generating the strong (pointwise) convergence of quantum channels. These continuity bounds are valid for arbitrary quantum channels (in contrast to the continuity bounds obtained in [9, 10]) and do not depend on channel characteristics at all.

2. Preliminaries

2.1. Basic notations and auxiliary lemmas

Let $\mathcal{H}$ be a finite-dimensional or separable infinite-dimensional Hilbert space. Denote by $\mathfrak{B}(\mathcal{H})$ the algebra of all bounded operators on $\mathcal{H}$ with the operator norm $\|\cdot\|$ and by $\mathfrak{T}(\mathcal{H})$ the Banach space of all trace-class operators on $\mathcal{H}$ with the trace norm $\|\cdot\|_1$. Let $\mathfrak{S}(\mathcal{H})$ be the set of quantum states (positive operators in $\mathfrak{T}(\mathcal{H})$ with unit trace) [6, 12].

The *Bures distance* between quantum states ρ and σ is defined as

$$\beta(\rho, \sigma) = \sqrt{2\left(1 - \sqrt{F(\rho, \sigma)}\right)}, \tag{1}$$

where $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$ is the *fidelity* of ρ and σ [6, 12]. The following relations between the Bures distance and the trace-norm distance hold

$$\tfrac{1}{2}\|\rho - \sigma\|_1 \leq \beta(\rho, \sigma) \leq \sqrt{\|\rho - \sigma\|_1}. \tag{2}$$

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

If quantum systems A and B are described by Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ then the composite quantum system AB is described by the Hilbert space $\mathcal{H}_{AB} \doteq \mathcal{H}_A \otimes \mathcal{H}_B$. For a state ρ_{AB} in $\mathfrak{S}(\mathcal{H}_{AB})$ denote by ρ_A and ρ_B its marginal states $\mathrm{Tr}_B \rho_{AB}$ and $\mathrm{Tr}_A \rho_{AB}$ correspondingly (here and in what follows Tr_X denotes the partial trace $\mathrm{Tr}_{\mathcal{H}_X}$ over the space $\mathcal{H}_X$).

The *von Neumann entropy* $H(\rho) = \mathrm{Tr}\eta(\rho)$ of a state $\rho \in \mathfrak{S}(\mathcal{H})$, where $\eta(x) = -x\log x$ if $x > 0$ and $\eta(0) = 0$, is a nonnegative concave lower semicontinuous function on the set $\mathfrak{S}(\mathcal{H})$, it is continuous on $\mathfrak{S}(\mathcal{H})$ if and only if $\dim \mathcal{H} < +\infty$ [6, 13].

The *quantum relative entropy* for states ρ and σ in $\mathfrak{S}(\mathcal{H})$ is defined as

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

where $\{|i\rangle\}$ is the orthonormal basis of eigenvectors of the state ρ and it is assumed that $H(\rho\|\sigma) = +\infty$ if $\mathrm{supp}\,\rho$ is not contained in $\mathrm{supp}\,\sigma$ [6, 13] (the support $\mathrm{supp}\,\rho$ of a positive operator ρ is the orthogonal complement to its kernel).

The *quantum mutual information* of a state ρ_{AB} of a bipartite quantum system is defined as

$$I(A\!:\!B)_\rho = H(\rho_{AB}\|\rho_A \otimes \rho_B) = H(\rho_A) + H(\rho_B) - H(\rho_{AB}),$$

where the second formula is valid if $H(\rho_{AB})$ is finite [14].

Properties of the quantum relative entropy imply that $\rho \mapsto I(A:B)_\rho$ is a nonnegative lower semicontinuous function on the set $\mathfrak{S}(\mathcal{H}_{AB})$. It is well known that

$$I(A\!:\!B)_\rho \leq 2 \min \{H(\rho_A), H(\rho_B)\} \tag{3}$$

for arbitrary state ρ_{AB} and that

$$I(A\!:\!B)_\rho \leq \min \{H(\rho_A), H(\rho_B)\} \tag{4}$$

for any separable state ρ_{AB} [14, 15].

The *quantum conditional mutual information (QCMI)* of a state ρ_{ABC} of a finite-dimensional tripartite quantum system is defined as

$$I(A\!:\!B|C)_\rho \doteq H(\rho_{AC}) + H(\rho_{BC}) - H(\rho_{ABC}) - H(\rho_C). \tag{5}$$

This quantity is essentially used in quantum information theory [6, 12].

To avoid possible uncertainty "$\infty - \infty$" in (5) in infinite dimensions one should define the QCMI for any state ρ_{ABC} by one of the equivalent expressions

$$I(A\!:\!B|C)_\rho = \sup_{P_A} \left[I(A\!:\!BC)_{Q_A\rho Q_A} - I(A\!:\!C)_{Q_A\rho Q_A} \right], \quad Q_A = P_A \otimes I_{BC}, \tag{6}$$

$$I(A\!:\!B|C)_\rho = \sup_{P_B} \left[I(B\!:\!AC)_{Q_B\rho Q_B} - I(B\!:\!C)_{Q_B\rho Q_B} \right], \quad Q_B = P_B \otimes I_{AC}, \tag{7}$$

where the suprema are taken over all finite rank projectors P_A in $\mathfrak{B}(\mathcal{H}_A)$ and P_B in $\mathfrak{B}(\mathcal{H}_B)$ [16]. We assume here that $I(X:Y)_{Q_X\rho Q_X} = cI(X:Y)_{c^{-1}Q_X\rho Q_X}$, where $c = \mathrm{Tr}\,Q_X\rho_{ABC}$.

Formulae (6) and (7) define the same lower semicontinuous function on the set $\mathfrak{S}(\mathcal{H}_{ABC})$ inheriting all basic properties of the finite-dimensional QCMI defined by (5) [16, Theorem 2]. We will use the following relation (chain rule)

$$I(X\!:\!YZ|C)_\rho = I(X\!:\!Y|C)_\rho + I(X\!:\!Z|YC)_\rho \tag{8}$$

valid for any state ρ in $\mathfrak{S}(\mathcal{H}_{XYZC})$ (with possible values $+\infty$ on both sides). Relation (8) can be proved by noting that it holds if the systems X, Y, Z and C are finite-dimensional and by using Corollary 9 in [16].

We will use the upper bound

$$I(A\!:\!B|C)_\rho \leq 2\min\{H(\rho_A), H(\rho_B), H(\rho_{AC}), H(\rho_{BC})\} \tag{9}$$

valid for any state ρ_{ABC}. It directly follows from upper bound (3) and the expression $I(X\!:\!Y|C)_\rho = I(X\!:\!YC)_\rho - I(X\!:\!C)_\rho$, $X, Y = A, B$, which is a partial case of (8).

The QCMI (defined in (6),(7)) satisfies the following relation

$$\left| pI(A\!:\!B|C)_\rho + (1-p)I(A\!:\!B|C)_\sigma - I(A\!:\!B|C)_{p\rho+(1-p)\sigma} \right| \leq h_2(p) \tag{10}$$

valid for $p \in (0,1)$ and any states ρ_{ABC}, σ_{ABC} with finite $I(A:B|C)_\rho$, $I(A:B|C)_\sigma$, where $h_2(p) = \eta(p) + \eta(1-p)$ is the binary entropy [17].

We will repeatedly use the following simple lemmas.

Lemma 1. *If U and V are isometries from $\mathcal{H}$ into $\mathcal{H}'$ then*

$$\|U\rho U^* - V\rho V^*\|_1 \leq 2\|(U-V)\rho\|_1 \leq 2\|U-V\| \quad \text{for any } \rho \in \mathfrak{S}(\mathcal{H}).$$

Lemma 2. [5] *If f is a concave nonnegative function on $[0, +\infty)$ then for any positive $x < y$ and any $z \geq 0$ the following inequality holds*

$$xf(z/x) \leq yf(z/y).$$

2.2. Set of quantum states with bounded energy

2.2.1. General case Let H_A be a positive (semi-definite) operator on a Hilbert space $\mathcal{H}_A$ treated as a Hamiltonian of a quantum system A. Let $E \geq E_0 \doteq \inf_{\|\varphi\|=1} \langle\varphi|H_A|\varphi\rangle$. Then

$$\mathfrak{C}_{H_A,E} = \{\rho \in \mathfrak{S}(\mathcal{H}_A) \,|\, \mathrm{Tr}H_A\rho \leq E\}$$

is a closed convex subset of $\mathfrak{S}(\mathcal{H}_A)$ consisting of states with the mean energy not exceeding E.

It is well known that the von Neumann entropy is continuous on the set $\mathfrak{C}_{H_A,E}$ for any $E > E_0$ if (and only if) the Hamiltonian H_A satisfies the condition

$$\mathrm{Tr}\, e^{-\lambda H_A} < +\infty \quad \text{for all } \lambda > 0 \tag{11}$$

and that the maximal value of the entropy on this set is achieved at the *Gibbs state* $\gamma_A(E) \doteq e^{-\lambda(E)H_A}/\mathrm{Tr}e^{-\lambda(E)H_A}$, where the parameter $\lambda(E)$ is determined by the equality $\mathrm{Tr}H_A e^{-\lambda(E)H_A} = E\mathrm{Tr}e^{-\lambda(E)H_A}$ [18]. Condition (11) implies that H_A is an unbounded operator having discrete spectrum of finite multiplicity [7].

We will use the function $F_{H_A}(E) \doteq \sup_{\rho \in \mathfrak{C}_{H_A,E}} H(\rho) = H(\gamma_A(E))$. It is easy to show that F_{H_A} is a strictly increasing concave function on $[E_0, +\infty)$ such that $F_{H_A}(E_0) = \log d_0$, where d_0 is the multiplicity of E_0 [5, 17].

Let

$$\bar{F}_{H_A}(E) = F_{H_A}(E + E_0) = H(\gamma_A(E + E_0)) \tag{12}$$

and $\bar{F}_{H_A}^{-1}$ be the inverse function to the function $\bar{F}_{H_A}$. The above-stated properties of F_{H_A} shows that $\bar{F}_{H_A}^{-1}$ is an increasing function on $[\log d_0, +\infty)$ taking values in $[0, +\infty)$.

Lemma 3. *Let* $\bar{E} \doteq E - E_0 \geq 0$, $d \geq d_0$ *and* $\gamma(d) \doteq \bar{F}_{H_A}^{-1}(\log d)$. *Let* B *be any system. If* $\bar{E} \leq \gamma(d)$ *then for any pure state* ρ *in* $\mathfrak{S}(\mathcal{H}_{AB})$ *such that* $\mathrm{Tr} H_A \rho_A \leq E$ *there is a pure state* σ *in* $\mathfrak{S}(\mathcal{H}_{AB})$ *such that* $\mathrm{rank}\,\sigma_A \leq d$, $\mathrm{Tr} H_A \sigma_A \leq E$, $\frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{\bar{E}/\gamma(d)}$ *and*

$$\|\rho - \sigma\|_1 \mathrm{Tr} \bar{H}_A \left[[\rho - \sigma]_-\right]_A \leq 2\bar{E}, \quad \|\rho - \sigma\|_1 \mathrm{Tr} \bar{H}_A \left[[\rho - \sigma]_+\right]_A \leq 2\bar{E},$$

where $\bar{H}_A = H_A - E_0 I_A$, $[\rho - \sigma]_-$ *and* $[\rho - \sigma]_+$ *are, respectively, the negative and positive parts of the Hermitian operator* $\rho - \sigma$.

Proof. Note first that

$$C \doteq \max_{1 \leq k \leq d} \{\langle k | \bar{H}_A | k \rangle\} \geq \gamma(d) \tag{13}$$

for any orthonormal set $\{|k\rangle\}_{k=1}^d \subset \mathcal{H}_A$. Indeed, if $P_d = \sum_{k=1}^d |k\rangle\langle k|$ then $\mathrm{Tr} P_d \bar{H}_A \leq Cd$. So, the entropy $\log d$ of the state $d^{-1} P_d$ does not exceed $F_{\bar{H}_A}(C) = \bar{F}_{H_A}(C)$.

Let $\rho = |\varphi\rangle\langle\varphi|$, where $|\varphi\rangle = \sum_{k=1}^{+\infty} \sqrt{p_k} |\alpha_k \otimes \beta_k\rangle$ for some base $\{|\alpha_k\rangle\}$ and $\{|\beta_k\rangle\}$ in the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. We may assume that the basis $\{|\alpha_k\rangle\}$ is reordered in such a way that the sequence $\{\langle \alpha_k | \bar{H}_A | \alpha_k \rangle\}_k$ is non-decreasing.

Let $|\psi\rangle = (1 - \delta_d)^{-1/2} \sum_{k=1}^d \sqrt{p_k} |\alpha_k \otimes \beta_k\rangle$, where $\delta_d = \sum_{k>d} p_k$, and $\sigma = |\psi\rangle\langle\psi|$. Since $\mathrm{Tr} \bar{H}_A \rho_A \leq \bar{E}$, we have $\delta_d \langle \alpha_d | \bar{H}_A | \alpha_d \rangle \leq \sum_{k>d} p_k \langle \alpha_k | \bar{H}_A | \alpha_k \rangle \leq \bar{E}$ and hence

$$\delta_d \leq \bar{E}/\langle \alpha_d | \bar{H}_A | \alpha_d \rangle \leq \bar{E}/\gamma(d) \leq 1,$$

where the second inequality follows from (13). It is easy to see that the above inequality implies $\mathrm{Tr} \bar{H}_A \sigma_A \leq \bar{E}$, i.e. $\mathrm{Tr} H_A \sigma_A \leq E$.

Since $\langle \varphi | \psi \rangle = (1 - \delta_d)^{1/2}$, we have $\frac{1}{2}\|\rho - \sigma\|_1 = \sqrt{1 - |\langle\varphi|\psi\rangle|^2} = \sqrt{\delta_d}$.

Assume that $\rho \neq \sigma$ (otherwise all the assertions are trivial). By using spectral decomposition of the operator $\rho - \sigma = |\varphi\rangle\langle\varphi| - |\psi\rangle\langle\psi|$ one can show that $[\rho - \sigma]_\pm$ are 1-rank operators corresponding to the vectors $|\gamma_\pm\rangle = p_\pm |\varphi\rangle + q_\pm |\psi\rangle$, where $p_\pm = \sqrt{(1 \pm x)/(2x)}$ and $q_\pm = -\sqrt{(1 \mp x)/2x}$, $x = \sqrt{\delta_d}$ (see the proof of Theorem 1 in [20]). So, we have

$$\mathrm{Tr} \bar{H}_A [[\rho - \sigma]_\pm]_A = \langle \gamma_\pm | \bar{H}_A \otimes I_B | \gamma_\pm \rangle = p_\pm^2 \langle \varphi | \bar{H}_A \otimes I_B | \varphi \rangle + q_\pm^2 \langle \psi | \bar{H}_A \otimes I_B | \psi \rangle$$

$$+ 2 p_\pm q_\pm \Re \langle \varphi | \bar{H}_A \otimes I_B | \psi \rangle = \left[p_\pm^2 + 2 p_\pm q_\pm (1 - \delta_d)^{-1/2} + q_\pm^2 (1 - \delta_d)^{-1}\right] S + p_\pm^2 R$$

$$= \frac{1}{2x} \left[\frac{x^2 S}{1 \pm x} + (1 \pm x) R\right] \leq \frac{1}{2x} \left[x^2 (1 \mp x) \bar{E} + (1 \pm x) \bar{E}\right] \leq \frac{\bar{E}}{x} = \frac{2\bar{E}}{\|\rho - \sigma\|_1},$$

where $S = \sum_{k \leq d} p_k \langle \alpha_k | \bar{H}_A | \alpha_k \rangle$ and $R = \sum_{k>d} p_k \langle \alpha_k | \bar{H}_A | \alpha_k \rangle$. In the last line we used the obvious inequalities $S \leq (1 - x^2)\bar{E}$, $R \leq \bar{E}$ and $x \leq 1$. $\square$

In this paper we essentially use the modification of the Alicki-Fannes-Winter method (cf.[2, 5, 19]) adapted for the set of states with bounded energy [20]. This modification makes it possible to prove uniform continuity of any locally almost affine

function‡ f on the set

$$\mathfrak{C}_{H_A,E}^{\mathrm{ext}} \doteq \{\rho \in \mathfrak{S}(\mathcal{H}_{AB}) \,|\, \rho_A \in \mathfrak{C}_{H_A,E}\} \qquad (B \text{ is any given system})$$

such that $|f(\rho_{AB})| \leq CH(\rho_A)$ for some $C \in \mathbb{R}_+$ provided that

$$F_{H_A}(E) = o(\sqrt{E}) \quad \text{as} \quad E \to +\infty. \tag{14}$$

By Lemma 1 in [20] condition (14) holds if and only if

$$\lim_{\lambda \to 0^+} \left[\mathrm{Tr}\, e^{-\lambda H_A}\right]^{\lambda} = 1. \tag{15}$$

Condition (15) is stronger than condition (11) (equivalent to $F_{H_A}(E) = o(E)$ as $E \to +\infty$) but the difference between these conditions is not too large. In terms of the sequence $\{E_k\}$ of eigenvalues of H_A condition (11) means that $\lim_{k\to\infty} E_k/\log k = +\infty$, while (15) is valid if $\liminf_{k\to\infty} E_k/\log^q k > 0$ for some $q > 2$ [20, Proposition 1].

It is essential that condition (15) holds for the Hamiltonian of a system of quantum oscillators considered in the next subsection.

2.2.2. The ℓ-mode quantum oscillator Let A be the ℓ-mode quantum oscillator with the frequencies $\omega_1, ..., \omega_\ell$. In this case

$$F_{H_A}(E) = \max_{\{x_i\}} \sum_{i=1}^{\ell} g(x_i/\hbar\omega_i - 1/2), \quad E \geq E_0 \doteq \frac{1}{2}\sum_{i=1}^{\ell} \hbar\omega_i,$$

where $g(x) = (x+1)\log(x+1) - x\log x$ and the maximum is over all ℓ-tuples $x_1,...,x_\ell$ such that $\sum_{i=1}^{\ell} x_i = E$ and $x_i \geq \frac{1}{2}\hbar\omega_i$ [6, Ch.12]. It is shown in [17] that

$$F_{H_A}(E) \leq F_{\ell,\omega}(E) \doteq \ell\log\frac{E+E_0}{\ell E_*} + \ell, \quad E_* = \left[\prod_{i=1}^{\ell}\hbar\omega_i\right]^{1/\ell}, \tag{16}$$

(throughout the paper log is the natural logarithm) and that upper bound (16) is ε-sharp for large E. So, the function

$$\bar{F}_{\ell,\omega}(E) \doteq F_{\ell,\omega}(E+E_0) = \ell\log\frac{E+2E_0}{\ell E_*} + \ell, \tag{17}$$

is an ε-sharp upper bound on the function $\bar{F}_{H_A}(E) \doteq F_{H_A}(E+E_0)$. By employing the function $\bar{F}_{\ell,\omega}$ one can define the sequence of positive numbers

$$\hat{\gamma}(d) \doteq \bar{F}_{\ell,\omega}^{-1}(\log d) = (\ell/e)E_*\sqrt[\ell]{d} - 2E_0, \quad d > e^{\bar{F}_{\ell,\omega}(0)}, \tag{18}$$

which can be used instead of the sequence $\gamma(d)$ in Lemma 3. It is clear that $\hat{\gamma}(d) \leq \gamma(d)$ for all d.

‡ This means that $|f(p\rho + (1-p)\sigma) - pf(\rho) - (1-p)f(\sigma)| \leq r(p) = o(1)$ as $p \to 0^+$ for any ρ and σ.

2.3. Some facts about ensembles of quantum states

2.3.1. Discrete ensembles A finite or countable set $\{\rho_i\}$ of quantum states and a corresponding probability distribution $\{p_i\}$ is usually called a *discrete ensemble* and denoted by $\{p_i, \rho_i\}$. The state $\bar{\rho} \doteq \sum_i p_i \rho_i$ is called the *average state* of this ensemble [6, 12].

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

$$\chi\left(\{p_i, \rho_i\}_{i=1}^m\right) \doteq \sum_{i=1}^m p_i H(\rho_i \| \bar{\rho}) = H(\bar{\rho}) - \sum_{i=1}^m p_i H(\rho_i),$$

where the second expression is valid if $H(\bar{\rho}) < +\infty$. This quantity is an upper bound for the classical information obtained from the ensemble by quantum measurements [21]. It is essentially used in analysis of information properties of quantum channels and systems [6, 12].

Assume that $\mathcal{H}_A = \mathcal{H}$ and that $\{|i\rangle\}_{i=1}^m$ is a basis in a Hilbert space $\mathcal{H}_B$. Then

$$\chi(\{p_i, \rho_i\}_{i=1}^m) = I(A\!:\!B)_{\hat{\rho}}, \text{ where } \hat{\rho}_{AB} = \sum_{i=1}^m p_i \rho_i \otimes |i\rangle\langle i|. \tag{19}$$

The state $\hat{\rho}_{AB}$ is called the *qc-state* determined by the ensemble $\{p_i, \rho_i\}_{i=1}^m$ [12].

In analysis of continuity of the Holevo quantity we will use three measures of divergence between ensembles $\mu = \{p_i, \rho_i\}$ and $\nu = \{q_i, \sigma_i\}$ described in detail in [22, 17].

The quantity

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

is a true metric on the set of all ensembles of quantum states considered as *ordered* collections of states with the corresponding probability distributions.

The main advantage of D_0 is a direct computability, but from the quantum information point of view it is natural 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 $\mathfrak{S}(\mathcal{H})$ (where $\delta(\rho)$ is the Dirac measure concentrated at a state ρ) rather than ordered (or disordered) collection of states. If we want to identify ensembles corresponding to the same probability measure then we have to use the factorization of D_0, i.e. the quantity

$$D_*(\mu, \nu) \doteq \inf_{\mu' \in \mathcal{E}(\mu), \nu' \in \mathcal{E}(\nu)} D_0(\mu', \nu') \tag{20}$$

as a measure of divergence between ensembles $\mu = \{p_i, \rho_i\}$ and $\nu = \{q_i, \sigma_i\}$, where $\mathcal{E}(\mu)$ and $\mathcal{E}(\nu)$ are the sets of all countable ensembles corresponding to the measures $\sum_i p_i \delta(\rho_i)$ and $\sum_i q_i \delta(\sigma_i)$ respectively.

It is shown in [17] that the factor-metric D_* coincides with the EHS-distance D_{ehs} between ensembles of quantum states proposed by Oreshkov and Calsamiglia in [22] and

that D_* induces the topology of weak convergence on the set of all ensembles (interpreted as probability measures).§ It is obvious that

$$D_*(\mu, \nu) \le D_0(\mu, \nu) \tag{21}$$

for any ensembles μ and ν.

We will also use the Kantorovich distance

$$D_K(\mu, \nu) = \frac{1}{2} \inf_{\{P_{ij}\}} \sum P_{ij} \|\rho_i - \sigma_j\|_1 \tag{22}$$

between ensembles $\mu = \{p_i, \rho_i\}$ and $\nu = \{q_i, \sigma_i\}$ of quantum states, where the infimum is taken over all joint probability distributions $\{P_{ij}\}$ such that $\sum_j P_{ij} = p_i$ for all i and $\sum_i P_{ij} = q_j$ for all j. Since $D_* = D_{\text{ehs}}$, it follows from the result in [22] that

$$D_*(\mu, \nu) \le D_K(\mu, \nu) \tag{23}$$

for any discrete ensembles μ and ν.

The Kantorovich distance has a natural extension to the set of all generalized (continuous) ensembles. It generates the topology of weak convergence on this set (see the next subsection).

If μ and ν are discrete ensembles of states in $\mathfrak{S}(\mathcal{H})$, where $d \doteq \dim \mathcal{H} < +\infty$, then Proposition 16 in [17] implies that

$$|\chi(\mu) - \chi(\nu)| \le \varepsilon \log d + 2g(\varepsilon), \tag{24}$$

where $\varepsilon = D_*(\mu, \nu)$ and $g(\varepsilon) = (1 + \varepsilon)h_2\left(\frac{\varepsilon}{1+\varepsilon}\right)$. Since $g(\varepsilon)$ is an increasing function, it follows from (21) and (23) that inequality (24) remains valid for $\varepsilon = D_0(\mu, \nu)$ and for $\varepsilon = D_K(\mu, \nu)$. Continuity bound (24) with $\varepsilon = D_K(\mu, \nu)$ is a refined version of the continuity bound for the Holevo quantity obtained in [22].

2.3.2.   Generalized (continuous) ensembles In the study of infinite-dimensional quantum systems the notion of *generalized (continuous) ensemble* defined as a Borel probability measure on the set of quantum states is widely used [6, 23]. We denote by $\mathcal{P}(\mathcal{H})$ the set of all Borel probability measures on $\mathfrak{S}(\mathcal{H})$ equipped with the topology of weak convergence [24, 25].‖ The set $\mathcal{P}(\mathcal{H})$ can be treated as a complete separable metric space [25]. It contains the dense subset $\mathcal{P}_0(\mathcal{H})$ of discrete measures (corresponding to discrete ensembles). The average state of a generalized ensemble $\mu \in \mathcal{P}(\mathcal{H})$ is defined as the barycenter of the measure μ, that is $\bar{\rho}(\mu) = \int_{\mathfrak{S}(\mathcal{H})} \rho\mu(d\rho)$.

For an ensemble $\mu \in \mathcal{P}(\mathcal{H}_A)$ its image $\Phi(\mu)$ under a quantum channel $\Phi : A \to B$ is defined as the ensemble in $\mathcal{P}(\mathcal{H}_B)$ corresponding to the measure $\mu \circ \Phi^{-1}$ on $\mathfrak{S}(\mathcal{H}_B)$, i.e. $\Phi(\mu)[\mathfrak{S}_B] = \mu[\Phi^{-1}(\mathfrak{S}_B)]$ for any Borel subset $\mathfrak{S}_B \subseteq \mathfrak{S}(\mathcal{H}_B)$, where $\Phi^{-1}(\mathfrak{S}_B)$ is the pre-image of $\mathfrak{S}_B$ under the map Φ. If $\mu = \{p_i, \rho_i\}$ then $\Phi(\mu) = \{p_i, \Phi(\rho_i)\}$.

§ This means that a sequence $\{\{p_i^n, \rho_i^n\}\}_n$ converges to an ensemble $\{p_i^0, \rho_i^0\}$ w.r.t. the metric D_* if and only if $\lim_{n\to\infty} \sum_i p_i^n f(\rho_i^n) = \sum_i p_i^0 f(\rho_i^0)$ for any continuous bounded function f on $\mathfrak{S}(\mathcal{H})$.
‖ A sequence $\{\mu_n\}$ of measures weakly converges to a measure μ_0 if $\lim_{n\to\infty} \int f(\rho)\mu_n(d\rho) = \int f(\rho)\mu_0(d\rho)$ for any continuous bounded function f on $\mathfrak{S}(\mathcal{H})$.

The Holevo quantity of a generalized ensemble $\mu \in \mathcal{P}(\mathcal{H})$ is defined as

$$\chi(\mu) = \int_{\mathfrak{S}(\mathcal{H})} H(\rho \,\|\, \bar{\rho}(\mu)) \mu(d\rho) = H(\bar{\rho}(\mu)) - \int_{\mathfrak{S}(\mathcal{H})} H(\rho)\mu(d\rho),$$

where the second formula is valid under the condition $H(\bar{\rho}(\mu)) < +\infty$ [23].

The Kantorovich distance (22) is extended to generalized ensembles μ and ν by the expression

$$D_K(\mu, \nu) = \frac{1}{2} \inf_{\Lambda \in \Pi(\mu,\nu)} \int_{\mathfrak{S}(\mathcal{H}) \times \mathfrak{S}(\mathcal{H})} \|\rho - \sigma\|_1 \Lambda(d\rho, d\sigma), \tag{25}$$

where $\Pi(\mu, \nu)$ is the set of all Borel probability measures on $\mathfrak{S}(\mathcal{H}) \times \mathfrak{S}(\mathcal{H})$ with the marginals μ and ν. By noting that $\frac{1}{2}\|\rho - \sigma\|_1 \leq 1$ for any states ρ and σ, we see that the Kantorovich distance (25) generates the weak convergence on the set $\mathcal{P}(\mathcal{H})$ [24].

3. Bures distance and Energy-Constrained Bures distance between quantum channels

A *quantum channel* Φ from a quantum system A to a quantum system B is a linear completely positive trace preserving map from $\mathfrak{T}(\mathcal{H}_A)$ into $\mathfrak{T}(\mathcal{H}_B)$ [6, 12].

For any quantum channel $\Phi : A \to B$ the Stinespring theorem implies existence of a Hilbert space $\mathcal{H}_E$ and an isometry $V_\Phi : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ such that

$$\Phi(\rho) = \mathrm{Tr}_E V_\Phi \rho V_\Phi^*, \quad \rho \in \mathfrak{T}(\mathcal{H}_A). \tag{26}$$

The quantum channel

$$\mathfrak{T}(\mathcal{H}_A) \ni \rho \mapsto \widehat{\Phi}(\rho) = \mathrm{Tr}_B V_\Phi \rho V_\Phi^* \in \mathfrak{T}(\mathcal{H}_E) \tag{27}$$

is called *complementary* to the channel Φ [6, Ch.6].

The set of quantum channels with finite input dimension $\dim \mathcal{H}_A$ is typically equipped with the metric induced by the *diamond norm*

$$\|\Phi\|_\diamond \doteq \sup_{\rho \in \mathfrak{S}(\mathcal{H}_{AR})} \|\Phi \otimes \mathrm{Id}_R(\rho)\|_1 \tag{28}$$

of a Hermitian-preserving superoperator $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$, where R is any system [26],[12, Ch.9]. This norm coincides with the norm of complete boundedness of the dual map $\Phi^* : \mathfrak{B}(\mathcal{H}_B) \to \mathfrak{B}(\mathcal{H}_A)$ to the map Φ [27].

For our purposes it is more convenient to use the equivalent metric on the set of quantum channels called the *Bures distance*. It is defined for given channels Φ and Ψ from A to B as

$$\beta(\Phi, \Psi) = \sup_{\rho \in \mathfrak{S}(\mathcal{H}_{AR})} \beta\left(\Phi \otimes \mathrm{Id}_R(\rho), \Psi \otimes \mathrm{Id}_R(\rho)\right), \tag{29}$$

where $\beta(\cdot, \cdot)$ in the r.h.s. is the Bures distance between quantum states defined in (1) and R is any system. This metric is related to the notion of *operational fidelity* for quantum channels introduced in [28]. It is studied in detail in [29, 30]. In particular, it is shown in [30] that the Bures distance (29) can be also defined as

$$\beta(\Phi, \Psi) = \inf \|V_\Phi - V_\Psi\|, \tag{30}$$

where the infimum is over all common Stinespring representations

$$\Phi(\rho) = \mathrm{Tr}_E V_\Phi \rho V_\Phi^* \quad \text{and} \quad \Psi(\rho) = \mathrm{Tr}_E V_\Psi \rho V_\Psi^*, \tag{31}$$

and that the infimum in (30) is attainable. Definition (29) and the relations (2) imply that

$$\tfrac{1}{2}\|\Phi - \Psi\|_\diamond \leq \beta(\Phi, \Psi) \leq \sqrt{\|\Phi - \Psi\|_\diamond}. \tag{32}$$

This shows the equivalence of the Bures distance and the diamond-norm distance on the set of all channels between given quantum systems.

In the case $\dim \mathcal{H}_A = +\infty$ the diamond-norm distance (the Bures distance) becomes singular: there are infinite-dimensional quantum channels with arbitrarily close physical parameters having the diamond-norm distance equal to 2 [10]. In this case it is natural to use the distance between quantum channels induced by the *energy-constrained diamond norm*

$$\|\Phi\|_\diamond^E \doteq \sup_{\rho \in \mathfrak{S}(\mathcal{H}_{AR}), \mathrm{Tr} H_A \rho_A \leq E} \|\Phi \otimes \mathrm{Id}_R(\rho)\|_1, \quad E > E_0, \tag{33}$$

of a Hermitian-preserving superoperator $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$, where H_A is the Hamiltonian of the input system A and $E_0 \doteq \inf_{\|\varphi\|=1} \langle\varphi|H_A|\varphi\rangle$ [9, 10] (slightly different energy-constrained diamond norms are used in [31]). If H_A is an unbounded operator having discrete spectrum of finite multiplicity (in particular, if it satisfies condition (11)) then any of the norms (33) generates the strong (pointwise) convergence of quantum channels, i.e. vanishing of $\|\Phi_n - \Phi_0\|_\diamond^E$ as $n \to +\infty$ is equivalent to

$$\lim_{n \to +\infty} \Phi_n(\rho) = \Phi_0(\rho) \quad \forall \rho \in \mathfrak{S}(\mathcal{H}_A) \tag{34}$$

for a sequence $\{\Phi_n\}$ of quantum channels from A to any system B [9, Proposition 3]. Note that the strong convergence is substantially weaker than the diamond norm convergence if the systems A and B are infinite-dimensional.

In this paper we will use the *energy-constrained Bures distance*

$$\beta_E(\Phi, \Psi) = \sup_{\rho \in \mathfrak{S}(\mathcal{H}_{AR}), \mathrm{Tr} H_A \rho_A \leq E} \beta(\Phi \otimes \mathrm{Id}_R(\rho), \Psi \otimes \mathrm{Id}_R(\rho)), \quad E > E_0, \tag{35}$$

between quantum channels Φ and Ψ from A to B, where R is any system.

Definitions (33),(35) and the relations (2) imply that

$$\tfrac{1}{2}\|\Phi - \Psi\|_\diamond^E \leq \beta_E(\Phi, \Psi) \leq \sqrt{\|\Phi - \Psi\|_\diamond^E}. \tag{36}$$

The properties of the energy-constrained Bures distance used below are collected in the following

Proposition 1. *Let H_A be any positive densely defined operator on $\mathcal{H}_A$. For any $E > E_0$ the function $(\Phi, \Psi) \mapsto \beta_E(\Phi, \Psi)$ defined in (35) is a real metric on the set of all quantum channels from A to B which can be represented as follows*

$$\beta_E(\Phi, \Psi) = \inf \sup_{\rho \in \mathfrak{S}(\mathcal{H}_A), \mathrm{Tr} H_A \rho \leq E} \sqrt{\mathrm{Tr}(V_\Phi - V_\Psi)\rho(V_\Phi^* - V_\Psi^*)}, \tag{37}$$

where the infimum is over all common Stinespring representations (31). For any given channels Φ and Ψ the following properties hold:

a) $\beta_E(\Phi, \Psi)$ *tends to* $\beta(\Phi, \Psi)$ *as* $E \to +\infty$;

b) *the function* $E \mapsto \beta_E(\Phi, \Psi)$ *is concave and nondecreasing on* $[E_0, +\infty)$;

c) *for any* $E \geq E_0$ *the infimum in (37) is attained at a pair* (V_Φ, V_Ψ) *of isometries forming representation (31).*

If the operator H_A *satisfies condition (11) then for any* $E > E_0$ *the metric* β_E *generates the strong (pointwise) convergence of quantum channels, i.e. vanishing of* $\beta_E(\Phi_n, \Phi_0)$ *as* $n \to +\infty$ *for a sequence* $\{\Phi_n\}$ *of quantum channels from* A *to any system* B *is equivalent to (34).*

Proof. The coincidence of (35) and (37) along with the property c) follow from the proof of Theorem 1 in [30]. For readers convenience we present a simplified version of these arguments for the case $\mathcal{A} = \mathfrak{B}(\mathcal{H})$ in the Appendix.

Property a) and the monotonicity of β_E are obvious. To prove the concavity take any $E_1, E_2 > E_0$. For arbitrary $\varepsilon > 0$ there exist states ρ_1 and ρ_2 in $\mathfrak{S}(\mathcal{H}_{AR})$ such that

$$\mathrm{Tr} H_A [\rho_i]_A \leq E_i \ \text{ and } \ \beta_{E_i}^2(\Phi, \Psi) \leq \beta^2(\Phi \otimes \mathrm{Id}_R(\rho_i), \Psi \otimes \mathrm{Id}_R(\rho_i)) + \varepsilon, \ \ i = 1, 2.$$

By invariance of the Bures distance (1) with respect to unitary transformation of both states we may assume that $\mathrm{supp} \mathrm{Tr}_A \rho_1 \perp \mathrm{supp} \mathrm{Tr}_A \rho_2$ and hence $\mathrm{supp} \Theta_1 \otimes \mathrm{Id}_R(\rho_1) \perp \mathrm{supp} \Theta_2 \otimes \mathrm{Id}_R(\rho_2)$ for any channels Θ_1 and Θ_2. Thus,

$$\beta_{\bar{E}}^2(\Phi, \Psi) \geq \beta^2(\Phi \otimes \mathrm{Id}_R(\bar\rho), \Psi \otimes \mathrm{Id}_R(\bar\rho)) = \tfrac{1}{2}\beta^2(\Phi \otimes \mathrm{Id}_R(\rho_1), \Psi \otimes \mathrm{Id}_R(\rho_1))$$

$$+ \tfrac{1}{2}\beta^2(\Phi \otimes \mathrm{Id}_R(\rho_2), \Psi \otimes \mathrm{Id}_R(\rho_2)) \geq \tfrac{1}{2}(\beta_{E_1}^2(\Phi, \Psi) + \beta_{E_2}^2(\Phi, \Psi)) - \varepsilon,$$

where $\bar\rho = \tfrac{1}{2}(\rho_1 + \rho_2)$ and $\bar{E} = \tfrac{1}{2}(E_1 + E_2)$. This implies the concavity of the function $E \mapsto \beta_{\bar{E}}^2(\Phi, \Psi)$. The concavity of the function $E \mapsto \beta_E(\Phi, \Psi)$ follows from the concavity and monotonicity of the function $\sqrt{x}$.

The last assertion of the proposition follows from (36) and Proposition 3 in [9], since condition (11) guarantees that H_A is an unbounded densely defined operator having discrete spectrum of finite multiplicity. $\square$

Corollary 1. *Let* Φ *and* Ψ *be arbitrary channels from* A *to* B *and* R *be any system. Let* H_A *be a positive operator on* $\mathcal{H}_A$ *and* β_E *the corresponding energy-constrained Bures distance defined in (35). There exists a common Stinespring representation (31) such that*

$$\|V_\Phi \otimes I_R \rho V_\Phi^* \otimes I_R - V_\Psi \otimes I_R \rho V_\Psi^* \otimes I_R\|_1 \leq 2\beta_E(\Phi, \Psi) \leq 2\sqrt{\|\Phi - \Psi\|_\diamond^E}$$

for any state ρ *in* $\mathfrak{S}(\mathcal{H}_{AR})$ *satisfying the inequality* $\mathrm{Tr} H_A \rho_A \leq E$.

Proof. Let $\widehat{V}_\Theta = V_\Theta \otimes I_R$, $\Theta = \Phi, \Psi$. By using Lemma 1 and the operator Cauchy-Schwarz inequality we obtain

$$\|\widehat{V}_\Phi \rho \widehat{V}_\Phi^* - \widehat{V}_\Psi \rho \widehat{V}_\Psi^*\|_1 \leq 2\|(\widehat{V}_\Phi - \widehat{V}_\Psi)\rho\|_1 = 2\mathrm{Tr} W(\widehat{V}_\Phi - \widehat{V}_\Psi)\rho$$

$$\leq 2\sqrt{\mathrm{Tr}(W(\widehat{V}_\Phi - \widehat{V}_\Psi)\rho(\widehat{V}_\Phi^* - \widehat{V}_\Psi^*)W^*)\mathrm{Tr}\rho} = 2\sqrt{\mathrm{Tr}(V_\Phi - V_\Psi)\rho_A(V_\Phi^* - V_\Psi^*)},$$

where W is the partial isometry from the polar decomposition of $(\widehat{V}_\Phi - \widehat{V}_\Psi)\rho$. So, the assertion of the corollary follows directly from Proposition 1 and (36).

By definition (27) the representations (30) and (37) imply the following

Corollary 2. *Let Φ and Ψ be arbitrary quantum channels from A to B.*

A) *There exist¶ complementary channels $\widehat{\Phi}$ and $\widehat{\Psi}$ from A to some system E such that $\beta(\widehat{\Phi}, \widehat{\Psi}) \leq \beta(\Phi, \Psi)$.*

B) *Let H_A be a positive operator on $\mathcal{H}_A$ and β_E the corresponding energy-constrained Bures distance defined in (35). There exist complementary channels $\widehat{\Phi}$ and $\widehat{\Psi}$ from A to some system E such that $\beta_E(\widehat{\Phi}, \widehat{\Psi}) \leq \beta_E(\Phi, \Psi)$.*

4. Basic lemma

Our main technical tool is the following lemma proved by modification of the Alicki-Fannes-Winter method [2, 5, 19, 20]. In this lemma we use the extended QCMI defined in (6),(7) and the function $g(x) = (1+x)h_2\left(\frac{x}{1+x}\right) = (x+1)\log(x+1) - x\log x$.

Lemma 4. *Let ρ and σ be any states in $\mathfrak{S}(\mathcal{H}_{ABCDR})$ such that $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon \leq 1$. Let $\mathcal{H}_*$ be a subspace of $\mathcal{H}_{AD}$ containing the supports of ρ_{AD} and σ_{AD}.*

A) *If $\dim \mathcal{H}_* = d < +\infty$ then $I(A{:}B|C)_\rho$ and $I(A{:}B|C)_\sigma$ are finite and*

$$|I(A{:}B|C)_\rho - I(A{:}B|C)_\sigma| \leq 2\varepsilon \log d + 2g(\varepsilon). \tag{38}$$

If ρ and σ are qc-states with respect to the decomposition $(AD)(BCR)$ then (38) holds with the term $2\varepsilon \log d$ replaced by $\varepsilon \log d$.+

B) *If $\mathrm{Tr} H_* \rho_{AD}, \mathrm{Tr} H_* \sigma_{AD} \leq E < +\infty$ for some positive operator H_* on $\mathcal{H}_*$ satisfying condition (15) then $I(A{:}B|C)_\rho$ and $I(A{:}B|C)_\sigma$ are finite and*

$$|I(A{:}B|C)_\rho - I(A{:}B|C)_\sigma| \leq 2\sqrt{2\varepsilon}F_{H_*}(E/\varepsilon) + 2g(\sqrt{2\varepsilon}), \tag{39}$$

where $F_{H_*}(E) \doteq \sup\{H(\rho)\,|\,\mathrm{supp}\rho \subseteq \mathcal{H}_*, \mathrm{Tr} H_*\rho \leq E\}$. If ρ and σ are pure states then (39) holds with ε replaced by $\varepsilon^2/2$.

C) *If $\rho_{BC} = \sigma_{BC}$ then factor 2 in the last terms in (38), (39) and in their specifications for qc-states and pure states can be omitted.*

Proof. A) In the proof of (38) we may assume that $\frac{1}{2}\|\rho - \sigma\|_1 = \varepsilon$. Following [5, 19] introduce the states $\gamma_+ = \varepsilon^{-1}[\rho - \sigma]_+$ and $\gamma_- = \varepsilon^{-1}[\rho - \sigma]_-$ in $\mathfrak{S}(\mathcal{H}_{ABCDR})$. Then

$$\frac{1}{1+\varepsilon}\rho + \frac{\varepsilon}{1+\varepsilon}\gamma_- = \omega^* = \frac{1}{1+\varepsilon}\sigma + \frac{\varepsilon}{1+\varepsilon}\gamma_+.$$

By taking partial trace we obtain

$$\frac{1}{1+\varepsilon}\rho_{ABC} + \frac{\varepsilon}{1+\varepsilon}[\gamma_-]_{ABC} = \omega^*_{ABC} = \frac{1}{1+\varepsilon}\sigma_{ABC} + \frac{\varepsilon}{1+\varepsilon}[\gamma_+]_{ABC}.$$

Basic properties of the QCMI and upper bound (3) imply

$$I(A{:}B|C)_\omega \leq I(A{:}BC)_\omega \leq I(AD{:}BCR)_\omega \leq 2H(\omega_{AD}) \tag{40}$$

¶ A complementary channel is defined up to the isometrical equivalence [32].
+ qc-states are defined in (19); we assume here that the basic $\{|i\rangle\}$ does not depend on a state.

for any state ω in $\mathfrak{S}(\mathcal{H}_{ABCDR})$. Since the operators $\mathrm{Tr}_{BCR}[\rho - \sigma]_{\pm}$ are supported on the subspace $\mathcal{H}_*$, it follows from (40) that

$$I(A\!:\!B|C)_\omega \leq 2\log d < +\infty, \quad \omega = \rho, \sigma, \gamma_+, \gamma_-. \tag{41}$$

By applying (10) to the above convex decompositions of ω^*_{ABC} we obtain

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

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

where $p = \frac{\varepsilon}{1+\varepsilon}$. By using (41) and nonnegativity of $I(A\!:\!B|C)$ we obtain (38).

If ρ and σ are qc-states with respect to the decomposition $(AD)(BCR)$ then the above states γ_+ and γ_- are qc-states as well. So, by using (4) instead of (3) one can remove factor 2 in (40) and (41).

B) We may consider $I(A\!:\!B|C)$ as a function on $\mathfrak{S}(\mathcal{H}_{BCR} \otimes \mathcal{H}_*)$. If $\varepsilon < \frac{1}{2}$ then the assertion can be proved by applying Proposition 1 in [20] to this function by using the nonnegativity of $I(A\!:\!B|C)$, upper bound (40) and inequality (10). If $\varepsilon \in [\frac{1}{2}, 1]$ then the required inequality follows from upper bound (40).

C) The possibility to remove factor 2 from the last term of (39) in the case $\rho_{BC} = \sigma_{BC}$ is shown in [33, Lemma 2]. By inserting the same arguments in the above proof of part A, it is easy to prove the similar assertion for (38). $\square$

5. Continuity bounds for the output conditional mutual information

The quantum mutual information and its conditional version play a basic role in analysis of informational properties of quantum systems and channels (see Sections 6-8). In this section we explore continuity properties of the QCMI (defined in (6),(7)) at the output of a channel acting on one subsystem of a tripartite system, i.e. the quantity $I(B:D|C)_{\Phi \otimes \mathrm{Id}_{CD}(\rho)}$, where $\Phi : A \to B$ is an arbitrary channel, C, D are any systems and ρ is a state in $\mathfrak{S}(\mathcal{H}_{ADC})$. We will obtain continuity bounds for the function

$$(\Phi, \rho) \mapsto I(B\!:\!D|C)_{\Phi \otimes \mathrm{Id}_{CD}(\rho)}$$

assuming that the set of all channels from A to B is equipped with the Bures distance (30) in the case $\dim \mathcal{H}_A < +\infty$ and with the energy-constrained Bures distance (35) in the case $\dim \mathcal{H}_A = +\infty$. We will also obtain continuity bounds for the function

$$\Phi \mapsto I(B^n\!:\!D|C)_{\Phi^{\otimes n} \otimes \mathrm{Id}_{CD}(\rho)}$$

for arbitrary $\rho \in \mathfrak{S}(\mathcal{H}_{A^nCD})$ and any natural n.

5.1. QCMI at the output of a local channel

5.1.1. Finite input dimension The following proposition contains continuity bounds for the function $(\Phi, \rho) \mapsto I(B:D|C)_{\Phi \otimes \mathrm{Id}_{CD}(\rho)}$, where Φ runs over all channels with a

finite-dimensional input system A and ρ runs over all states of a tripartite system ACD.

Proposition 2. *Let Φ and Ψ be quantum channels from a finite-dimensional system A to arbitrary system B, C and D be any systems, $d_A \doteq \dim \mathcal{H}_A$. Then*

$$|I(B\!:\!D|C)_{\Phi\otimes\mathrm{Id}_{CD}(\rho)} - I(B\!:\!D|C)_{\Psi\otimes\mathrm{Id}_{CD}(\sigma)}| \leq 2\varepsilon \log d_A + 2\varepsilon \log 2 + 2g(\varepsilon), \quad (42)$$

for any states ρ and σ in $\mathfrak{S}(\mathcal{H}_{ACD})$, where $\varepsilon = \frac{1}{2}\|\rho - \sigma\|_1 + \beta(\Phi, \Psi)$.

If $\Phi = \Psi$ then the term $2\varepsilon \log 2$ in (42) can be omitted. If $\rho = \sigma$ then (42) holds without factor 2 in the last term.

Continuity bound (42) is tight in both cases $\Phi = \Psi$ and $\rho = \sigma$ for any system C. The Bures distance $\beta(\Phi, \Psi)$ in (42) can be replaced by $\sqrt{\|\Phi - \Psi\|_\diamond}$.

Proof. By Theorem 1 in [30] there is a common Stinespring representation (31) such that $\|V_\Phi - V_\Psi\| = \beta(\Phi, \Psi)$. Then $\hat{\rho} = V_\Phi \otimes I_{CD} \rho V_\Phi^* \otimes I_{CD}$ and $\hat{\sigma} = V_\Psi \otimes I_{CD} \sigma V_\Psi^* \otimes I_{CD}$ are extensions of the states $\Phi \otimes \mathrm{Id}_{CD}(\rho)$ and $\Psi \otimes \mathrm{Id}_{CD}(\sigma)$. Lemma 1 implies

$$\|\hat{\rho} - \hat{\sigma}\|_1 \leq \|\rho - \sigma\|_1 + 2\|V_\Phi - V_\Psi\|. \quad (43)$$

Since the states $\hat{\rho}_{BE} = V_\Phi \rho_A V_\Phi^*$ and $\hat{\sigma}_{BE} = V_\Psi \sigma_A V_\Psi^*$ are supported on the subspace $V_\Phi \mathcal{H}_A \vee V_\Psi \mathcal{H}_A$ of $\mathcal{H}_{BE}$ whose dimension does not exceed $2d_A$, Lemma 4A and inequality (43) imply (42). If $\Phi = \Psi$ then the above states $\hat{\rho}_{BE}$ and $\hat{\sigma}_{BE}$ are supported on the d_A-dimensional subspace $V_\Phi \mathcal{H}_A = V_\Psi \mathcal{H}_A$. If $\rho = \sigma$ then $\hat{\rho}_{CD} = \hat{\sigma}_{CD}$ and (42) holds with $2g(\varepsilon)$ replaced by $g(\varepsilon)$ by Lemma 4C.

The tightness of continuity bound (42) in the case $\Phi = \Psi$ follows from Corollary 2 in [17].

The tightness of continuity bound (42) in the case $\rho = \sigma$ can be derived from the tightness of continuity bound (81) for the quantum capacity. It can be also shown directly by using the erasure channels $\Phi_{1/2}$ and $\Phi_{1/2-x}$ (see the proof of Theorem 1 in Section 7).

The last assertion of the proposition follows from the right inequality in (32) and monotonicity of the function $g(x)$. $\square$

5.1.2. Finite input energy Note first that the (uniform) continuity bound for the function $(\Phi, \rho) \mapsto I(B\!:\!D|C)_{\Phi\otimes\mathrm{Id}_{CD}(\rho)}$ under the energy constraint on ρ_A can be obtained by combining (uniform) continuity bounds for the functions $\rho \mapsto I(B:D|C)_{\Phi\otimes\mathrm{Id}_{CD}(\rho)}$ and $\Phi \mapsto I(B:D|C)_{\Phi\otimes\mathrm{Id}_{CD}(\rho)}$ under this constraint *not depending* on Φ and ρ.

Proposition 3. *Let $\Phi : A \to B$ be an arbitrary quantum channel, C and D any systems. If the Hamiltonian H_A of the input system A satisfies condition (15) then the function $\rho_{ACD} \mapsto I(B\!:\!D|C)_{\Phi\otimes\mathrm{Id}_{CD}(\rho)}$ is uniformly continuous on the set of states with bounded energy of ρ_A. Quantitatively,*

$$|I(B\!:\!D|C)_{\Phi\otimes\mathrm{Id}_C(\rho)} - I(B\!:\!D|C)_{\Phi\otimes\mathrm{Id}_C(\sigma)}| \leq 2\sqrt{2\varepsilon}\bar{F}_{H_A}(\bar{E}/\varepsilon) + 2g(\sqrt{2\varepsilon}) \quad (44)$$

for any states ρ and σ in $\mathfrak{S}(\mathcal{H}_{ACD})$ such that $\mathrm{Tr}H_A\rho_A, \mathrm{Tr}H_A\sigma_A \leq E$ and $\frac{1}{2}\|\rho-\sigma\|_1 \leq \varepsilon$, where $\bar{F}_{H_A}$ is the function defined in (12) and $\bar{E} = E - E_0$.

If ρ and σ are pure states then (44) holds with ε replaced by $\varepsilon^2/2$.

If A is the ℓ-mode quantum oscillator then the function $\bar{F}_{H_A}$ in (44) can be replaced by its upper bound $\bar{F}_{\ell,\omega}$ defined in (17).

The main term in (44) tends to zero as $\varepsilon \to 0^+$, since condition (15) implies that $\bar{F}_{H_A}(E) = o(\sqrt{E})$ as $E \to +\infty$ (see Section 2.2.1).

Proof. Assume that the channel Φ has the Stinespring representation (26). Continuity bound (44) follows from Lemma 4B with $\mathcal{H}_* = V_\Phi \mathcal{H}_A \subseteq \mathcal{H}_{BE}$ and $H_* = V_\Phi H_A V_\Phi^* - E_0 I_{\mathcal{H}_*}$. $\square$

The continuity bound for the function $\Phi \mapsto I(B:D|C)_{\Phi \otimes \mathrm{Id}_{CD}(\rho)}$ under the energy constraint on ρ_A not depending on Φ and ρ is presented in Proposition 5 below (the case $n = 1$).

5.2. QCMI at the output of n copies of a local channel

5.2.1. Finite input dimension The following proposition contains continuity bound for the function $\Phi \mapsto I(B^n:D|C)_{\Phi^{\otimes n} \otimes \mathrm{Id}_{CD}(\rho)}$, where Φ runs over all channels with a finite-dimensional input system A and ρ is a fixed state of a tripartite system ACD. It is proved by using the Leung-Smith telescopic method from [1].

Proposition 4. *Let Φ and Ψ be quantum channels from a finite-dimensional system A to arbitrary system B, C and D be any systems and $n \in \mathbb{N}$. Then*

$$\left| I(B^n:D|C)_{\Phi^{\otimes n} \otimes \mathrm{Id}_{CD}(\rho)} - I(B^n:D|C)_{\Psi^{\otimes n} \otimes \mathrm{Id}_{CD}(\rho)} \right| \leq n(2\varepsilon \log(2d_A) + g(\varepsilon)) \qquad (45)$$

for any state ρ in $\mathfrak{S}(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_{CD})$, where $\varepsilon = \beta(\Phi, \Psi)$ and $d_A = \dim \mathcal{H}_A$.

Continuity bound (45) is tight for any system C (for each given n and large d_A). The Bures distance $\beta(\Phi, \Psi)$ in (45) can be replaced by $\sqrt{\|\Phi - \Psi\|_\diamond}$.

Proof. Consider the states $\sigma_k = \Phi^{\otimes k} \otimes \Psi^{\otimes(n-k)} \otimes \mathrm{Id}_{CD}(\rho)$, $k = 0, 1, ..., n$. Then

$$\left| I(B^n:D|C)_{\sigma_n} - I(B^n:D|C)_{\sigma_0} \right| = \left| \sum_{k=1}^{n} I(B^n:D|C)_{\sigma_k} - I(B^n:D|C)_{\sigma_{k-1}} \right|$$
$$\leq \sum_{k=1}^{n} \left| I(B^n:D|C)_{\sigma_k} - I(B^n:D|C)_{\sigma_{k-1}} \right|. \qquad (46)$$

For each k the chain rule (8) implies that

$$I(B^n:D|C)_{\sigma_k} - I(B^n:D|C)_{\sigma_{k-1}} = I(Y_k:D|C)_{\sigma_k} + I(B_k:D|Y_kC)_{\sigma_k}$$

$$- I(Y_k:D|C)_{\sigma_{k-1}} - I(B_k:D|Y_kC)_{\sigma_{k-1}} = I(B_k:D|Y_kC)_{\sigma_k} - I(B_k:D|Y_kC)_{\sigma_{k-1}}, \qquad (47)$$

where $Y_k = B_1..B_{k-1}B_{k+1}..B_n$. The second equality here follows from the equality $\mathrm{Tr}_{B_k}\sigma_k = \mathrm{Tr}_{B_k}\sigma_{k-1}$. Note that the finite entropy of the states $\rho_{A_1}, ..., \rho_{A_n}$, upper bound (9) and monotonicity of the QCMI under local channels guarantee finiteness of all the terms in (46) and (47).

By Theorem 1 in [30] there is a common Stinespring representation (31) such that $\|V_\Phi - V_\Psi\| = \beta(\Phi, \Psi)$. To estimate the last difference in (47) consider the states

$$\hat{\sigma}_k = W_k \otimes V_\Phi^k \otimes I_{CD} \cdot \rho \cdot W_k^* \otimes [V_\Phi^k]^* \otimes I_{CD}, \qquad (48)$$

$$\hat{\sigma}_{k-1} = W_k \otimes V_\Psi^k \otimes I_{CD} \cdot \rho \cdot W_k^* \otimes [V_\Psi^k]^* \otimes I_{CD} \tag{49}$$

in $\mathfrak{S}(\mathcal{H}_{B^n E^n CD})$, where $W_k = V_\Phi^1 \otimes \ldots \otimes V_\Phi^{k-1} \otimes V_\Psi^{k+1} \otimes \ldots \otimes V_\Psi^n$ is an isometry from $\mathcal{H}_{A^n \setminus A_k}$ into $\mathcal{H}_{[BE]^n \setminus [BE]_k}$, $V_\Phi^k \cong V_\Phi$ and $V_\Psi^k \cong V_\Psi$ are isometries from $\mathcal{H}_{A_k}$ into $\mathcal{H}_{B_k E_k}$. It follows from (31) that these states are extensions of the states σ_k and σ_{k-1}, i.e. $\mathrm{Tr}_{E^n} \hat{\sigma}_k = \sigma_k$ and $\mathrm{Tr}_{E^n} \hat{\sigma}_{k-1} = \sigma_{k-1}$. Since $[\hat{\sigma}_k]_{B_k E_k} = V_\Phi^k \rho_{A_k} [V_\Phi^k]^*$, $[\hat{\sigma}_{k-1}]_{B_k E_k} = V_\Psi^k \rho_{A_k} [V_\Psi^k]^*$ and $\mathrm{Tr}_{B_k E_k} \hat{\sigma}_k = \mathrm{Tr}_{B_k E_k} \hat{\sigma}_{k-1}$, Lemma 4C implies that

$$|I(B_k : D | Y_k C)_{\sigma_k} - I(B_k : D | Y_k C)_{\sigma_{k-1}}| \leq 2\varepsilon' \log(2 d_A) + g(\varepsilon'), \tag{50}$$

where $\varepsilon' = \frac{1}{2}\|\hat{\sigma}_k - \hat{\sigma}_{k-1}\|_1$. By using Lemma 1 we obtain

$$\varepsilon' \leq \|W_k \otimes V_\Phi^k \otimes I_{CD} - W_k \otimes V_\Psi^k \otimes I_{CD}\| = \|V_\Phi - V_\Psi\| = \beta(\Phi, \Psi) = \varepsilon.$$

Hence, it follows from (47) and (50) that $|I(B^n : D | C)_{\sigma_k} - I(B^n : D | C)_{\sigma_{k-1}}|$ does not exceed $2\varepsilon \log(2 d_A) + g(\varepsilon)$. So, (46) implies the required inequality.

The tightness of the continuity bound in Proposition 4 follows from the tightness of continuity bound (42) in the case $\rho = \sigma$, since for arbitrary channel $\Phi : A \to B$, any system D and a state $\rho \in \mathfrak{S}(\mathcal{H}_{AD})$ we have $I(B^n : D^n)_{\Phi^{\otimes n} \otimes \mathrm{Id}_{D^n}(\rho^{\otimes n})} = n I(B : D)_{\Phi \otimes \mathrm{Id}_D(\rho)}$.

The last assertion of the proposition follows from the right inequality in (32) and monotonicity of the function $g(x)$. $\square$

5.2.2. Finite input energy In this subsection we assume that H_A is the Hamiltonian of a system A satisfying condition (15). We will use the function $\bar{F}_{H_A}$ defined in (12) and the increasing sequence $\{\gamma(d) = \bar{F}_{H_A}^{-1}(\log d)\}_{d \geq d_0}$ introduced in Lemma 3 tending to $+\infty$ as $d \to +\infty$, where d_0 is the multiplicity of the minimal eigenvalue E_0 of H_A. We will also use the energy-constrained Bures distance β_E between quantum channels and the energy-constrained diamond norm $\| \cdot \|_\diamond^E$ defined, respectively, in (35) and (33).

Proposition 5. *Let Φ and Ψ be arbitrary quantum channels from A to B, C and D any systems, $n \in \mathbb{N}$ and ρ a state in $\mathfrak{S}(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_{CD})$ such that $\sum_{k=1}^n \mathrm{Tr} H_A \rho_{A_k} \leq nE$. Let $s = 0$ if the function $E \mapsto \bar{F}_{H_A}(E)/\sqrt{E}$ is non-increasing on $\mathbb{R}_+$ and $s = 1$ otherwise. Let $t = 0$ if $\mathrm{Tr} H_A \rho_{A_k} \leq E$ for all k and $t = 1$ otherwise. Then*

$$\left| I(B^n : D | C)_{\Phi^{\otimes n} \otimes \mathrm{Id}_{CD}(\rho)} - I(B^n : D | C)_{\Psi^{\otimes n} \otimes \mathrm{Id}_{CD}(\rho)} \right| \leq n(T_{s,t}(E, \varepsilon) + g(\varepsilon) + 2\varepsilon \log 2), \tag{51}$$

where $\varepsilon = \beta_E(\Phi, \Psi)$ and

$$T_{s,t}(E, \varepsilon) \doteq \min_{\gamma(d) \geq 2\bar{E}} \left[\left(4 \sqrt{\frac{2^s \bar{E}}{\gamma(d)}} + \frac{4 s t \bar{E}}{\gamma(d)} + 2\varepsilon \right) \log d + 4 g \left(\sqrt{\frac{2^s \bar{E}}{\gamma(d)}} \right) \right]. \tag{52}$$

(the minimum here is over natural numbers d such that $\frac{1}{2}\gamma(d) \geq \bar{E} \doteq E - E_0$).

Inequality (51) also holds with $\varepsilon = \sqrt{\|\Phi - \Psi\|_\diamond^E}$.

For given $E \geq E_0$, s and t the quantity $T_{s,t}(E, \varepsilon)$ tends to zero as $\varepsilon \to 0^+$.

Remark 1. The quantity $T_{s,t}(E, \varepsilon)$ is determined by the function $\bar{F}_{H_A}$ (since $\gamma(d)$ is determined by $\bar{F}_{H_A}$), i.e by the Hamiltonian H_A of the system A. It can be calculated for any $\varepsilon > 0$ and $E > E_0$ by using explicit expression for $\bar{F}_{H_A}$. If this expression is

not known (or too complicated) we can use any upper bound $\widehat{F}_{H_A}$ for $\bar{F}_{H_A}$ provided that $\widehat{F}_{H_A}$ is a concave nonnegative function on $[0, +\infty)$ such that $\widehat{F}_{H_A}(E) = o(\sqrt{E})$ as $E \to +\infty$. It follows from the proofs of Proposition 5 and Lemma 3 that *the quantity $T_{s,t}(E, \varepsilon)$ in Proposition 5 can be determined by formula (52) with $\gamma(d)$ and s replaced, respectively, by $\hat{\gamma}(d) \doteq \widehat{F}_{H_A}^{-1}(\log d)$ and the corresponding $\hat{s}$.* This approach will be used to specify Proposition 5 for the case when A is a multi-mode quantum oscillator (see Corollary 3 below).

Proof of Proposition 5. We can repeat the arguments from the proof of Proposition 4 up to the estimation of the last difference in (47). In the case $st = 1$ we have to estimate this difference by two different ways depending on the value of $x_k \doteq \mathrm{Tr}H_A\rho_{A_k}$.

Let d be a natural number such that $\gamma(d) \geq 2\bar{E}$, N_1 the set of all indexes k such that $\bar{x}_k \doteq x_k - E_0 \leq \gamma(d)/2$ and $N_2 = \{1, ..., n\} \setminus N_1$. Let $n_i = \sharp(N_i)$, $X_i = n_i^{-1}\sum_{k \in N_i} x_k$ and $\bar{X}_i = X_i - E_0$, $i = 1, 2$. Let $\hat{\rho}$ be a pure state in $\mathfrak{S}(H_{A^nCDR})$ such that $\mathrm{Tr}_R\hat{\rho} = \rho$.

Assume first that $k \in N_1$. Then $\Delta_d^k \doteq \sqrt{2\bar{x}_k/\gamma(d)} \leq 1$. Lemma 3 implies existence of a pure state ϱ^k in $\mathfrak{S}(\mathcal{H}_{A^nCDR})$ such that $\mathrm{rank}\varrho_{A_k}^k \leq d$, $\mathrm{Tr}H_A\varrho_{A_k}^k \leq x_k$, $\frac{1}{2}\|\hat{\rho} - \varrho^k\|_1 \leq \sqrt{\bar{x}_k/\gamma(d)}$ and

$$\|\hat{\rho} - \varrho^k\|_1 \mathrm{Tr}\bar{H}_A[[\hat{\rho} - \varrho^k]_{\pm}]_{A_k} \leq 2\bar{x}_k, \quad \bar{H}_A = H_A - E_0 I_A. \tag{53}$$

By Corollary 1 for each k there exists a common Stinespring representation (31) with the isometries V_Φ^k and V_Ψ^k from $\mathcal{H}_A$ into $\mathcal{H}_{BE_k}$ such that

$$\tfrac{1}{2}\|V_\Phi^k \otimes I_{T_k}\, \varrho^k\, [V_\Phi^k]^* \otimes I_{T_k} - V_\Psi^k \otimes I_{T_k}\, \varrho^k\, [V_\Psi^k]^* \otimes I_{T_k}\|_1 \leq \varepsilon_k \doteq \beta_{x_k}(\Phi, \Psi),$$

where $T_k = A^nCDR \setminus A_k$. Consider the pure states

$$\hat{\varsigma}_k^k = W_k \otimes V_\Phi^k \otimes I_{CDR} \cdot \varrho^k \cdot W_k^* \otimes [V_\Phi^k]^* \otimes I_{CDR}, \tag{54}$$

$$\hat{\varsigma}_{k-1}^k = W_k \otimes V_\Psi^k \otimes I_{CDR} \cdot \varrho^k \cdot W_k^* \otimes [V_\Psi^k]^* \otimes I_{CDR} \tag{55}$$

in $\mathfrak{S}(\mathcal{H}_{B^nE_1...E_nCDR})$, where W_k is the same isometry as in (48),(49), while V_Φ^k, V_Ψ^k are the isometries chosen before. Then $\|\hat{\varsigma}_k^k - \hat{\varsigma}_{k-1}^k\|_1 \leq \varepsilon_k$. Since $\mathrm{rank}\varrho_{A_k}^k \leq d$, Lemma 4C implies that

$$|I(B_k:D|Y_kC)_{\hat{\varsigma}_k^k} - I(B_k:D|Y_kC)_{\hat{\varsigma}_{k-1}^k}| \leq 2\varepsilon_k \log(2d) + g(\varepsilon_k) \qquad (Y_k = B^n \setminus B_k).$$

Let $\hat{\sigma}_k$ and $\hat{\sigma}_{k-1}$ be the pure states in $\mathfrak{S}(\mathcal{H}_{B^nE_1...E_nCDR})$ defined, respectively, by formulas (54) and (55) with ϱ^k replaced by $\hat{\rho}$. Since $\mathrm{Tr}H_A\varrho_{A_k}^k, \mathrm{Tr}H_A\rho_{A_k} \leq x_k$ and $\|\hat{\sigma}_{k-1} - \hat{\varsigma}_{k-1}^k\|_1 = \|\hat{\sigma}_k - \hat{\varsigma}_k^k\|_1 = \|\hat{\rho} - \varrho^k\|_1 \leq 2\Delta_d^k \leq 2$, Lemma 4B with $\mathcal{H}_* = V_\Theta^k\mathcal{H}_A \subseteq \mathcal{H}_{B_kE_k}$ and $H_* = V_\Theta^kH_A[V_\Theta^k]^* - E_0 I_{\mathcal{H}_*}$, $\Theta = \Phi, \Psi$, implies that

$$|I(B_k:D|Y_kC)_{\sigma_{k-1}} - I(B_k:D|Y_kC)_{\hat{\varsigma}_{k-1}^k}| \text{ and } |I(B_k:D|Y_kC)_{\sigma_k} - I(B_k:D|Y_kC)_{\hat{\varsigma}_k^k}| \tag{56}$$

are bounded above by

$$2\Delta_d^k\bar{F}_{H_A}\left(\frac{2\bar{x}_k}{(\Delta_d^k)^2}\right) + 2g(\Delta_d^k) = 2\sqrt{\frac{2\bar{x}_k}{\gamma(d)}}\log d + 2g\left(\sqrt{\frac{2\bar{x}_k}{\gamma(d)}}\right), \tag{57}$$

where the equality $\bar{F}_{H_A}(\gamma(d)) = \log d$ was used.

Thus, the last difference in (47) is bounded above by

$$B_k^1(d) \doteq 4\sqrt{\frac{2\bar{x}_k}{\gamma(d)}}\log d + 4g\left(\sqrt{\frac{2\bar{x}_k}{\gamma(d)}}\right) + 2\varepsilon_k\log(2d) + g(\varepsilon_k).$$

Let $\bar{\varepsilon} = n_1^{-1}\sum_{k\in N_1}\varepsilon_k$. By using the concavity of the functions $\sqrt{x}$ and $g(x)$ along with the monotonicity of $g(x)$, we obtain

$$\sum_{k\in N_1} B_k^1(d) \leq 4n_1\sqrt{\frac{2\bar{X}_1}{\gamma(d)}}\log d + 4g\left(\sqrt{\frac{2\bar{X}_1}{\gamma(d)}}\right) + 2\bar{\varepsilon}n_1\log(2d) + n_1 g(\bar{\varepsilon}), \qquad (58)$$

For each $k \in N_2$ let $\hat{\sigma}_k$ and $\hat{\sigma}_{k-1}$ be the states in $\mathfrak{S}(\mathcal{H}_{B^n E_1...E_n CD})$ defined, respectively, by formulas (48) and (49) with arbitrary sets $\{V_\Phi^j\}$ and $\{V_\Psi^j\}$ of isometries from common Stinespring representation (31). Since $I(B_k:D|Y_kC) \leq I(B_kE_k:D|Y_kC)$, upper bound (9) implies that in this case the last difference in (47) is bounded above by

$$B_k^2 \doteq \max\{I(B_kE_k:D|Y_kC)_{\hat{\sigma}_k}, I(B_kE_k:D|Y_kC)_{\hat{\sigma}_{k-1}}\} \leq 2H(\rho_{A_k}) \leq 2F_{H_A}(x_k).$$

Since $(n - n_2)X_1 + n_2X_2 \leq nE$ and $X_1 \geq E_0$, we have $X_2 \leq n\bar{E}/n_2 + E_0$. So, by using concavity and monotonicity of the function F_{H_A} we obtain

$$\sum_{k\in N_2} B_k^2 \leq 2\sum_{k\in N_2} F_{H_A}(x_k) \leq 2n_2 F_{H_A}(X_2) \leq 2n_2\bar{F}_{H_A}(n\bar{E}/n_2). \qquad (59)$$

It is easy to see that $X_1 \leq E$. So, since the function $E \mapsto \beta_E(\Phi, \Psi)$ is concave and nondecreasing by Proposition 1, we have

$$\bar{\varepsilon} = \frac{1}{n_1}\sum_{k\in N_1}\beta_{x_k}(\Phi, \Psi) \leq \beta_{X_1}(\Phi, \Psi) \leq \beta_E(\Phi, \Psi) = \varepsilon. \qquad (60)$$

Since $\bar{x}_k > \gamma(d)/2$ for all $k \in N_2$ and $(n - n_2)E_0 + \sum_{k\in N_2}\bar{x}_k + n_2 E_0 \leq \sum_{k\in N_1} x_k + \sum_{k\in N_2} x_k \leq nE$, we have $n_2/n \leq 2\bar{E}/\gamma(d)$. So, it follows from (46),(47),(58)-(60) and Lemma 2 that the l.h.s. of (51) is bounded above by

$$\sum_{k\in N_1} B_k^1(d) + \sum_{k\in N_2} B_k^2 \leq n\left(4\sqrt{\frac{2\bar{E}}{\gamma(d)}} + 2\varepsilon\right)\log d + 4ng\left(\sqrt{\frac{2\bar{E}}{\gamma(d)}}\right)$$
$$+ 2n\varepsilon\log 2 + ng(\varepsilon) + \frac{4n\bar{E}}{\gamma(d)}\bar{F}_{H_A}\left(\frac{\gamma(d)}{2}\right). \qquad (61)$$

Since $\bar{F}_{H_A}(\gamma(d)/2) \leq \bar{F}_{H_A}(\gamma(d)) = \log d$, it implies (51).

If $t = 0$, i.e. $\text{Tr}H_A\rho_{A_k} \leq E$ for all k then the above-defined set N_2 is empty for all d such that $\gamma(d) \geq 2\bar{E}$. So, the last term in (61) disappears in this case.

If $s = 0$, i.e. the function $E \mapsto \bar{F}_{H_A}(E)/\sqrt{E}$ is non-increasing then we can make more subtle estimate of the quantities in (56) by using property (53).

Let $\alpha_k = \frac{1}{2}\|\hat{\rho} - \varrho^k\|_1$. Consider the states $\gamma_-^k = \alpha_k^{-1}[\hat{\sigma}_k - \hat{\varsigma}_k^k]_-$ and $\gamma_+^k = \alpha_k^{-1}[\hat{\sigma}_k - \hat{\varsigma}_k^k]_+$ which are determined by formula (54) with ϱ^k replaced, respectively, by $\alpha_k^{-1}[\hat{\rho} - \varrho^k]_-$ and $\alpha_k^{-1}[\hat{\rho} - \varrho^k]_+$. We have

$$\frac{1}{1+\alpha_k}\hat{\sigma}_k + \frac{\alpha_k}{1+\alpha_k}\gamma_-^k = \omega^* = \frac{1}{1+\alpha_k}\hat{\varsigma}_k^k + \frac{\alpha_k}{1+\alpha_k}\gamma_+^k.$$

By using the arguments from the proof of Lemma 4A (based on relation (10)) we obtain that the quantity $\left| I(B_k:D|Y_kC)_{\hat{\sigma}_k} - I(B_k:D|Y_kC)_{\hat{\varsigma}_k^k} \right|$ is bounded above by

$$\alpha_k \max\left\{ I(B_k:D|Y_kC)_{\gamma_-^k}, I(B_k:D|Y_kC)_{\gamma_+^k} \right\} + 2g(\alpha_k) \tag{62}$$

Basic properties of the QCMI and upper bound (9) imply

$$I(B_k:D|Y_kC)_{\gamma_\pm^k} \leq I(B_kE_k:D|Y_kC)_{\gamma_\pm^k} \leq 2H([\gamma_\pm^k]_{B_kE_k}) \leq 2H\left(\alpha_k^{-1}[[\hat{\rho} - \varrho^k]_\pm]_{A_k}\right).$$

It follows from (53) that $H(\alpha_k^{-1}[[\hat{\rho} - \varrho^k]_\pm]_{A_k}) \leq F_{\bar{H}_A}(\bar{x}_k/\alpha_k^2) = \bar{F}_{H_A}(\bar{x}_k/\alpha_k^2)$. So, (62) is bounded above by

$$2\alpha_k \bar{F}_{H_A}\left(\frac{\bar{x}_k}{\alpha_k^2}\right) + 2g(\alpha_k) \leq 2\sqrt{\frac{\bar{x}_k}{\gamma(d)}}\bar{F}_{H_A}(\gamma(d)) + 2g\left(\sqrt{\frac{\bar{x}_k}{\gamma(d)}}\right)$$

$$= 2\sqrt{\frac{\bar{x}_k}{\gamma(d)}}\log d + 2g\left(\sqrt{\frac{\bar{x}_k}{\gamma(d)}}\right). \tag{63}$$

Since $\alpha_k \leq \sqrt{\bar{x}_k/\gamma(d)}$, the inequality in (63) follows from the assumed non-increasing property of the function $E \mapsto \bar{F}_{H_A}(E)/\sqrt{E}$ and the non-decreasing property of g. The equality follows from the definition of $\gamma(d)$.

Similar arguments show that the quantity $\left| I(B_k:D|Y_kC)_{\hat{\sigma}_{k-1}} - I(B_k:D|Y_kC)_{\hat{\varsigma}_{k-1}^k} \right|$ is also bounded above by (63).

If $\bar{x}_k > \gamma(d)$ then $\sqrt{\bar{x}_k/\gamma(d)}\bar{F}_{H_A}(\gamma(d)) \geq \bar{F}_{H_A}(\bar{x}_k) = F_{H_A}(x_k)$. So, (9) implies that the r.h.s. of (63) is an upper bound for the quantities in (56) for all k (regardless of the value of x_k). So, the splitting into the sets N_1 and N_2 is not needed in this case.

The validity of inequality (51) with $\varepsilon = \sqrt{\|\Phi - \Psi\|_\diamond^E}$ follows from (36).

The last assertion of the proposition can be easily proved by noting that $\bar{F}_{H_A}(E) = o(\sqrt{E})$ as $E \to +\infty$ by condition (15). $\square$

Corollary 3. *Let A be the ℓ-mode quantum oscillator with the frequencies $\omega_1, ..., \omega_\ell$ and the assumptions of Proposition 5 hold. Let $F_{\ell,\omega}(E) = \ell[\log((E + E_0)/\ell E_*) + 1]$ be a ε-sharp upper bound for $F_{H_A}(E)$, where $E_0 \doteq \frac{1}{2}\sum_{i=1}^\ell \hbar\omega_i$ and $E_* \doteq \left[\prod_{i=1}^\ell \hbar\omega_i\right]^{1/\ell}$. Then*

$$\left| I(B^n:D|C)_{\Phi^{\otimes n}\otimes \mathrm{Id}_{CD}(\rho)} - I(B^n:D|C)_{\Psi^{\otimes n}\otimes \mathrm{Id}_{CD}(\rho)} \right| \leq n(P_r(E,\varepsilon) + g(\varepsilon) + 2\varepsilon\log 2), \tag{64}$$

for any $r \in (0,1]$ provided that $\varepsilon = \beta_E(\Phi, \Psi) \leq 1/r$, where

$$P_r(E,\varepsilon) = 2\varepsilon(1 + 2r)F_{\ell,\omega}(E) + 4\ell(2 + 1/r)\eta(\varepsilon r) + 4g(\varepsilon r) + 6\varepsilon e^{-\ell}, \tag{65}$$

$\eta(x) = -x\log x$. Continuity bound (64) with optimal r is tight for large E and any given n. Continuity bound (64) also holds with $\varepsilon = \sqrt{\|\Phi - \Psi\|_\diamond^E} \leq 1/r$.

Remark 2. The parameter r in (64) can be used to optimize this continuity bound for given E and ε. By choosing $r \sim 1/F_{\ell,\omega}(E)$ and by noting that $F_{\ell,\omega}(E) = F_{H_A}(E) + o(1)$ for large E, we see that the r.h.s. of (64) can be made not greater than

$$2\varepsilon n F_{\ell,\omega}(E) + o(F_{\ell,\omega}(E)) = 2\varepsilon n F_{H_A}(E) + o(F_{H_A}(E)) \quad \text{as} \quad E \to +\infty.$$

Proof. By Remark 1 the continuity bound (64) can be derived from Proposition 5 by using the function $\bar{F}_{\ell,\omega}$ defined in (17) as an upper bound for $\bar{F}_{H_A}$ and the corresponding sequence $\hat{\gamma}(d)$ defined in (18). Lemma 5 below implies that the function $E \mapsto \bar{F}_{\ell,\omega}(E)/\sqrt{E}$ is non-increasing. So, we may set $s = 0$ in formula (52).

Let $\delta \in (0, 1]$ and d_δ the minimal natural number such that $\bar{E}/\hat{\gamma}(d_\delta) \leq \delta^2$. By using the definition of d_δ and (18) we obtain

$$\log d_\delta \leq \log \left(\left[\frac{\bar{E}\delta^{-2} + 2E_0}{\ell E_* e^{-1}} \right]^\ell + 1 \right) \leq \ell \log \left[\frac{\bar{E}\delta^{-2} + 2E_0}{\ell E_* e^{-1}} \right] + \left[\frac{\ell E_* e^{-1}}{\bar{E}\delta^{-2} + 2E_0} \right]^\ell$$

$$\leq \ell \log \left[\frac{E + E_0}{\ell E_* \delta^2 e^{-1}} \right] + e^{-\ell} = F_{\ell,\omega}(E) - 2\ell \log \delta + e^{-\ell},$$

where the second inequality follows from the inequality $\log(1 + x) \leq x$ and the third inequality holds, since $\delta \leq 1$ and $2E_0 \geq E_*$. Thus, by setting $\delta = r\varepsilon$ and using monotonicity of the function $g(x)$ it is easy to show that $T_{0,t}(E, \varepsilon)$ is bounded above by the quantity $P_r(E, \varepsilon)$ defined in (65).

The tightness of continuity bound (64) with optimal r follows from the tightness of continuity bound (95) for the quantum capacity. It can be also shown directly by using the erasure channels $\Phi_{1/2}$ and $\Phi_{1/2-x}$ (see the proof of Theorem 2 in Section 8). $\square$

Lemma 5. *The function* $f(x) = x \log(a/x^2 + b)$ *is increasing on* $(0, +\infty)$ *provided that* $a > 0$ *and* $b \geq e/2$.

Proof. This assertion is proved by calculation of the derivative $f'(x)$ followed by a simple estimation.

6. Continuity bounds for the output Holevo quantity

In analysis of information properties of quantum channels we have to consider the output Holevo quantity of a given channel $\Phi : A \to B$ corresponding to a discrete or continuous ensemble μ of input quantum states, i.e. the quantity

$$\chi(\Phi(\mu)) = \int H(\Phi(\rho) \| \Phi(\bar{\rho}(\mu))) \mu(d\rho) = H(\Phi(\bar{\rho}(\mu))) - \int H(\Phi(\rho)) \mu(d\rho),$$

where the second formula is valid under the condition $H(\Phi(\bar{\rho}(\mu))) < +\infty$.

We will consider the output Holevo quantity $\chi(\Phi(\mu))$ as a function of a pair (channel Φ, input ensemble μ) assuming that

- the set of discrete ensembles is equipped with one of the metrics D_0, D_* and D_K described in Section 2.3.1;

- the set of generalized (continuous) ensembles is equipped with the Kantorovich metric D_K defined in (25);

- the set of quantum channels is equipped with the Bures distance β defined in (29) in the case $\dim \mathcal{H}_A < +\infty$ and with the energy-constrained Bures distance β_E defined in (35) in the case $\dim \mathcal{H}_A = +\infty$.

6.1. Finite input dimension

Speaking about the output Holevo quantity $\chi(\Phi(\mu))$ of a channel Φ with finite-dimensional input space we restrict attention to discrete ensembles μ, i.e. $\mu = \{p_i, \rho_i\}$, for which

$$\chi(\Phi(\mu)) = \chi(\{p_i, \Phi(\rho_i)\}) \doteq \sum_i p_i H(\Phi(\rho_i) \| \Phi(\bar\rho)), \quad \bar\rho = \sum_i p_i \rho_i.$$

Tight continuity bound for the function $(\Phi, \mu) \mapsto \chi(\Phi(\mu))$ depending on the input dimension of a channel is presented in the following proposition.

Proposition 6. *Let Φ and Ψ be quantum channels from a finite-dimensional system A to arbitrary system B. Let μ and ν be discrete ensembles of states in $\mathfrak{S}(\mathcal{H}_A)$. Then*

$$|\chi(\Phi(\mu)) - \chi(\Psi(\nu))| \leq \varepsilon \log d_A + \varepsilon \log 2 + 2g(\varepsilon), \tag{66}$$

where $d_A \doteq \dim \mathcal{H}_A$, $\varepsilon = D_*(\mu, \nu) + \beta(\Phi, \Psi)$ and $g(\varepsilon) = (1+\varepsilon)h_2\left(\frac{\varepsilon}{1+\varepsilon}\right)$.

If $\Phi = \Psi$ then the term $\varepsilon \log 2$ in (66) can be removed. If $\mu = \nu$ then (66) holds without factor 2 in the last term.

Continuity bound (66) is tight in both cases $\Phi = \Psi$ and $\mu = \nu$. The metric D_* in (66) can be replaced by any of the metrics D_0 and D_K, the Bures distance $\beta(\Phi, \Psi)$ can be replaced by $\sqrt{\|\Phi - \Psi\|_\diamond}$.

Proof. Assume that $\mu = \{p_i, \rho_i\}$ and $\nu = \{q_i, \sigma_i\}$. Take any $\epsilon > 0$. Let $\{\tilde{p}_i, \tilde{\rho}_i\}$ and $\{\tilde{q}_i, \tilde{\sigma}_i\}$ be ensembles belonging respectively to the sets $\mathcal{E}(\{p_i, \rho_i\})$ and $\mathcal{E}(\{q_i, \sigma_i\})$ such that

$$D_*(\{p_i, \rho_i\}, \{q_i, \sigma_i\}) \geq D_0(\{\tilde{p}_i, \tilde{\rho}_i\}, \{\tilde{q}_i, \tilde{\sigma}_i\}) - \epsilon \tag{67}$$

(see the definition (20) of D_*). By Theorem 1 in [30] there is a common Stinespring representation (31) such that $\|V_\Phi - V_\Psi\| = \beta(\Phi, \Psi)$. Consider the qc-states

$$\hat\rho = \sum_i \tilde{p}_i V_\Phi \tilde{\rho}_i V_\Phi^* \otimes |i\rangle\langle i| \quad \text{and} \quad \hat\sigma = \sum_i \tilde{q}_i V_\Psi \tilde{\sigma}_i V_\Psi^* \otimes |i\rangle\langle i|$$

in $\mathfrak{S}(\mathcal{H}_{BED})$, where $\{|i\rangle\}$ is a basic in $\mathcal{H}_D$. Representation (19) implies that

$$\chi(\{p_i, \Phi(\rho_i)\}) = \chi(\{\tilde{p}_i, \Phi(\tilde{\rho}_i)\}) = I(B:D)_{\hat\rho},$$

$$\chi(\{q_i, \Psi(\sigma_i)\}) = \chi(\{\tilde{q}_i, \Psi(\tilde{\sigma}_i)\}) = I(B:D)_{\hat\sigma},$$

where the first equalities follow from the obvious observation:

$$\tilde{\mu} \in \mathcal{E}(\mu) \quad \Rightarrow \quad \Theta(\tilde{\mu}) \in \mathcal{E}(\Theta(\mu)) \quad \Rightarrow \quad \chi(\Theta(\tilde{\mu})) = \chi(\Theta(\mu))$$

valid for any ensemble μ and any channel Θ.

Lemma 1 implies that

$$\|\hat{\rho} - \hat{\sigma}\|_1 \leq \sum_i \|\tilde{p}_i \tilde{\rho}_i - \tilde{q}_i \tilde{\sigma}_i\|_1 + 2\|V_\Phi - V_\Psi\| \leq 2(D_*(\mu, \nu) + \beta(\Phi, \Psi) + \epsilon). \qquad (68)$$

Since the states $\hat{\rho}_{BE}$ and $\hat{\sigma}_{BE}$ are supported on the subspace $V_\Phi \mathcal{H}_A \vee V_\Psi \mathcal{H}_A$ of $\mathcal{H}_{BE}$ whose dimension does not exceed $2d_A$, Lemma 4A and (67),(68) imply (66). If $\Phi = \Psi$ then the above states $\hat{\rho}_{BE}$ and $\hat{\sigma}_{BE}$ are supported on the d_A-dimensional subspace $V_\Phi \mathcal{H}_A = V_\Psi \mathcal{H}_A$. If $\mu = \nu$ then $\hat{\rho}_D = \hat{\sigma}_D$ and hence (66) holds with the term $2g(\varepsilon)$ replaced by $g(\varepsilon)$ by Lemma 4C.

The tightness of continuity bound (66) in the case $\Phi = \Psi$ follows from the tightness of the continuity bound (24), see Proposition 16 in [17].

The tightness of continuity bound (66) in the case $\mu = \nu$ follows from the tightness of continuity bound (79) for the Holevo capacity (which is derived from (66)). It can be also shown directly by using the erasure channels $\Phi_{1/2}$ and $\Phi_{1/2-x}$ (see the proof of Theorem 1 in Section 7).

Since the function $g(x)$ is increasing, the last assertion of the proposition follows from inequalities (21),(23) and the right inequality in (32). $\square$

6.2. Finite input energy

Speaking about the output Holevo quantity $\chi(\Phi(\mu))$ of a channel Φ between infinite-dimensional quantum systems A and B we will assume that μ runs over the set $\mathcal{P}(\mathcal{H}_A)$ of all generalized (continuous) ensembles equipped with the Kantorovich metric D_K defined in (25) generating the topology of weak convergence on this set. Specifications concerning the case of discrete ensembles will be given as additional remarks.

We will analyse the function $(\Phi, \mu) \mapsto \chi(\Phi(\mu))$ under the constraint on the average energy of μ, i.e. under the condition

$$E(\mu) \doteq \mathrm{Tr} H_A \bar{\rho}(\mu) = \int \mathrm{Tr} H_A \rho \mu(d\rho) \leq E,$$

where H_A is the Hamiltonian of the system A and $E \geq E_0 \doteq \inf_{\|\varphi\|=1} \langle \varphi | H_A | \varphi \rangle$.

Continuity bound for the function $(\Phi, \mu) \mapsto \chi(\Phi(\mu))$ under the constraint on the average energy of μ can be obtained by combining continuity bounds for the functions $\mu \mapsto \chi(\Phi(\mu))$ and $\Phi \mapsto \chi(\Phi(\mu))$ *not depending* on Φ and μ presented in the following two propositions.

Proposition 7. *Let* $\Phi : A \to B$ *be an arbitrary quantum channel. If the Hamiltonian* H_A *of the input system* A *satisfies condition (15) then the function*

$\mu \mapsto \chi(\Phi(\mu))$ *is uniformly continuous on the subset of* $\mathcal{P}(\mathcal{H}_A)$ *consisting of ensembles* μ *with bounded average energy* $E(\mu) \doteq \mathrm{Tr} H_A \bar\rho(\mu)$. *Quantitatively,*

$$|\chi(\Phi(\mu)) - \chi(\Phi(\nu))| \le 2\sqrt{2\varepsilon}\bar{F}_{H_A}(\bar{E}/\varepsilon) + 2g(\sqrt{2\varepsilon}) \tag{69}$$

for any ensembles μ *and* ν *such that* $E(\mu), E(\nu) \le E$ *and* $D_K(\mu,\nu) \le \varepsilon \le 1$, *where* $\bar{E} = E - E_0$ *and* $\bar{F}_{H_A}$ *is the function defined in* (12).

If μ *and* ν *are discrete ensembles then the Kantorovich metric* D_K *can be replaced by any of the metrics* D_0 *and* D_* *(described in Section 2.3.1).*

If A *is the* ℓ*-mode quantum oscillator then the function* $\bar{F}_{H_A}$ *in* (69) *can be replaced by its upper bound* $\bar{F}_{\ell,\omega}$ *defined in* (17).

Proof. For discrete ensembles μ and ν the inequality (69) with ε in $[D_*(\mu,\nu), 1]$ is proved in [20, Proposition 7]. It follows from (21) and (23) that this inequality holds for any $\varepsilon \in [D(\mu,\nu), 1]$, where D is either D_0 or D_K.

Let μ and ν be arbitrary generalized ensembles. The construction from the proof of Lemma 1 in [23] and the lower semicontinuity of the function $\mu \mapsto \chi(\Phi(\mu))$ ([23, Proposition 1]) allow to obtain sequences $\{\mu_n\}$ and $\{\nu_n\}$ of discrete ensembles weakly converging, respectively, to μ and ν such that

$$\lim_{n\to\infty} \chi(\Phi(\mu_n)) = \chi(\Phi(\mu)), \quad \lim_{n\to\infty} \chi(\Phi(\nu_n)) = \chi(\Phi(\nu)) \tag{70}$$

and $\bar\rho(\mu_n) = \bar\rho(\mu)$, $\bar\rho(\nu_n) = \bar\rho(\nu)$ for all n. Since inequality (69) holds for the ensembles μ_n and ν_n for all n and $D_K(\mu_n, \nu_n)$ tends to $D_K(\mu, \nu)$ as $n \to +\infty$, relations (70) imply the validity of (69) for the ensembles μ and ν. $\square$

Proposition 8. *Let* μ *be an arbitrary ensemble in* $\mathcal{P}(\mathcal{H}_A)$ *such that* $E(\mu) \doteq \mathrm{Tr} H_A \bar\rho(\mu) \le E$. *If the Hamiltonian* H_A *of the input system* A *satisfies condition* (15) *then the function* $\Phi \mapsto \chi(\Phi(\mu))$ *is uniformly continuous on the set of all quantum channels from* A *to any system* B *with respect to the strong convergence* (34). *Quantitatively,*

$$|\chi(\Phi(\mu)) - \chi(\Psi(\mu))| \le T_{s,0}(E, \varepsilon) + g(\varepsilon) + 2\varepsilon \log 2 \tag{71}$$

for any channels Φ *and* Ψ, *where* $\varepsilon = \beta_E(\Phi, \Psi)$ *and* $T_{s,0}(E, \varepsilon)$ *is the quantity defined in Proposition 5. Inequality* (71) *also holds with* $\varepsilon = \sqrt{\|\Phi - \Psi\|_\diamond^E}$.

If A *is the* ℓ*-mode quantum oscillator then* (71) *holds with* $T_{s,0}(E, \varepsilon)$ *replaced by the quantity* $P_r(E, \varepsilon)$ *defined in* (65) *for any* $r \in (0, 1]$ *such that* $\varepsilon \le 1/r$, *where* $\varepsilon = \beta_E(\Phi, \Psi)$ *or* $\varepsilon = \sqrt{\|\Phi - \Psi\|_\diamond^E}$.

Proof. Continuity bound (71) implies the first assertion of the proposition by the last assertions of Propositions 1 and 5.

If μ is a discrete ensemble then inequality (71) is derived from Proposition 5 with $n = 1$ and trivial C by using representation (19).

If μ is an arbitrary ensemble then the validity of (71) can be proved by approximation (by the same way as in the proof of Proposition 7).

The last assertion of the proposition follows from Corollary 3. $\square$

7. Continuity bounds for basic capacities of channels with finite input dimension.

Continuity bounds for basic capacities of quantum channels with finite output dimension d_B are obtained by Leung and Smith in [1]. The main term in all these bounds has the form $C\varepsilon \log d_B$, where C is a constant and ε is a distance between two channels (the diamond norm of their difference). These continuity bounds are essentially refined in [17] by applying the modification of the Leung-Smith approach (consisting in using the quantum conditional mutual information instead of the conditional entropy).

In this section we consider quantum channels with finite input dimension d_A and obtain continuity bounds for basic capacities of such channels with the main term $C\varepsilon \log d_A$, where ε is the Bures distance between quantum channels described in Section 3.

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

$$C_\chi(\Phi) = \sup_{\{p_i, \rho_i\}} \chi(\{p_i, \Phi(\rho_i)\}), \tag{72}$$

where the supremum is over all discrete ensembles of input states. This quantity characterizes the ultimate rate of classical information transmission through a channel provided that nonentangled input encoding is used [6, 12].

By the Holevo-Schumacher-Westmoreland theorem (cf.[34, 35]) the *classical capacity* of a quantum channel $\Phi : A \to B$ is given by the regularized expression

$$C(\Phi) = \lim_{n \to +\infty} n^{-1} C_\chi(\Phi^{\otimes n}). \tag{73}$$

If the weak additivity property holds for the channel Φ then $C(\Phi) = C_\chi(\Phi)$ [6, 36, 37].

The *entanglement-assisted classical capacity* of a quantum channel characterizes the ultimate rate of classical information transmission provided that an entangled state between the input and the output of a channel is used as an additional resource (see details in [6, 12]). By the Bennett-Shor-Smolin-Thaplyal theorem (cf. [38]) the entanglement-assisted classical capacity of a channel $\Phi : A \to B$ is given by the expression

$$C_{\mathrm{ea}}(\Phi) = \sup_{\rho \in \mathfrak{S}(\mathcal{H}_A)} I(\Phi, \rho), \tag{74}$$

in which

$$I(\Phi, \rho) = I(B\!:\!R)_{\Phi \otimes \mathrm{Id}_R(\hat{\rho})}, \tag{75}$$

where $\mathcal{H}_R \cong \mathcal{H}_A$ and $\hat{\rho}$ is a pure state in $\mathfrak{S}(\mathcal{H}_{AR})$ such that $\hat{\rho}_A = \rho$. The quantity $I(\Phi, \rho)$ is called the quantum mutual information of a channel Φ at a state ρ [6].

The *quantum capacity* of a channel characterizes the ultimate rate of quantum information transmission through a channel (see details in [6, 12]). By the Lloyd-Devetak-Shor theorem (cf. [39, 40]) the quantum capacity of a channel $\Phi : A \to B$ is given by the regularized expression

$$Q(\Phi) = \lim_{n \to +\infty} n^{-1} \bar{Q}(\Phi^{\otimes n}), \tag{76}$$

in which $\bar{Q}(\Phi)$ denotes the maximal value of the quantum coherent information $I_c(\Phi, \rho) \doteq H(\Phi(\rho)) - H(\widehat{\Phi}(\rho))$ over all input states $\rho \in \mathfrak{S}(\mathcal{H}_A)$ (where $\widehat{\Phi}$ is the complementary channel to the channel Φ defined in (27)).

The *private capacity* is the capacity of a quantum channel for classical communication with the additional requirement that almost no information is sent to the environment (see details in [6, 12]). By the Devetak theorem (cf. [39]) the private capacity of a channel $\Phi : A \to B$ is given by the regularized expression

$$C_{\mathrm{p}}(\Phi) = \lim_{n \to +\infty} n^{-1} \bar{C}_{\mathrm{p}}(\Phi^{\otimes n}), \tag{77}$$

where

$$\bar{C}_{\mathrm{p}}(\Phi) = \sup_{\{p_i, \rho_i\}} \left[\chi(\{p_i, \Phi(\rho_i)\}) - \chi(\{p_i, \widehat{\Phi}(\rho_i)\}) \right] \tag{78}$$

(the supremum is over all discrete ensembles of input states and $\widehat{\Phi}$ is the complementary channel to the channel Φ defined in (27)).

Now we consider continuity bounds for all the above capacities depending on the input dimension d_A. For the entanglement-assisted classical capacity the tight continuity bound

$$|C_{\mathrm{ea}}(\Phi) - C_{\mathrm{ea}}(\Psi)| \leq 2\varepsilon \log d_A + g(\varepsilon),$$

where $\varepsilon = \frac{1}{2}\|\Phi - \Psi\|_{\diamond}$, is obtained in [17, Proposition 29]. It is easy to show (by using formula (84) below with $n = 1$) that the same continuity bound holds for the quantity $\bar{Q}(\Phi)$ coinciding with $Q(\Phi)$ for degradable channels Φ [41].

For others basic capacities tight and close-to-tight continuity bounds depending on input dimension are presented in the following theorem, in which the Bures distance $\beta(\Phi, \Psi)$ described in Section 3 is used as a measure of divergence between channels Φ and Ψ and its is assumed that expressions (72)-(78) remain valid in the case $\dim \mathcal{H}_B = +\infty$.

Theorem 1. *Let Φ and Ψ be quantum channels from a finite-dimensional system A to arbitrary system B. Then*

$$|C_{\chi}(\Phi) - C_{\chi}(\Psi)| \leq \varepsilon \log d_A + \varepsilon \log 2 + g(\varepsilon), \tag{79}$$

$$|C(\Phi) - C(\Psi)| \leq 2\varepsilon \log d_A + 2\varepsilon \log 2 + g(\varepsilon), \tag{80}$$

$$|Q(\Phi) - Q(\Psi)| \leq 2\varepsilon \log d_A + 2\varepsilon \log 2 + g(\varepsilon), \tag{81}$$

$$|\bar{C}_{\mathrm{p}}(\Phi) - \bar{C}_{\mathrm{p}}(\Psi)| \leq 2\varepsilon \log d_A + 2\varepsilon \log 2 + 2g(\varepsilon), \tag{82}$$

$$|C_{\mathrm{p}}(\Phi) - C_{\mathrm{p}}(\Psi)| \leq 4\varepsilon \log d_A + 4\varepsilon \log 2 + 2g(\varepsilon), \tag{83}$$

where $d_A \doteq \dim \mathcal{H}_A$, $\varepsilon = \beta(\Phi, \Psi)$ and $g(\varepsilon) = (1 + \varepsilon)h_2\left(\frac{\varepsilon}{1+\varepsilon}\right)$.

The continuity bounds (79),(81) and (82) are tight, other continuity bounds are close-to-tight (up to factor 2 in the main term). The Bures distance $\beta(\Phi, \Psi)$ in (79)-(83) can be replaced by $\sqrt{\|\Phi - \Psi\|_{\diamond}}$.

Proof. Continuity bound (79) directly follows from Proposition 6 in Section 6.1 and the definition of the Holevo capacity.

Continuity bound (80) is obtained by using representation (19), Proposition 4 in Section 5.2 and Lemma 12 in [1].

Continuity bound (81) is proved by representing the coherent information as follows $I_c(\Phi, \rho) = I(B:R)_{\Phi \otimes \mathrm{Id}_R}(\hat{\rho}) - H(\rho)$, where $\hat{\rho}$ is a purification in $\mathfrak{S}(\mathcal{H}_{AR})$ of a state ρ. For arbitrary quantum channels Φ and Ψ, any n and a state ρ in $\mathfrak{S}(\mathcal{H}_{A^n})$ this implies

$$I_c(\Phi^{\otimes n}, \rho) - I_c(\Psi^{\otimes n}, \rho) = I(B^n:R)_{\Phi^{\otimes n} \otimes \mathrm{Id}_R}(\hat{\rho}) - I(B^n:R)_{\Psi^{\otimes n} \otimes \mathrm{Id}_R}(\hat{\rho}), \tag{84}$$

where $\hat{\rho}$ is a purification in $\mathfrak{S}(\mathcal{H}_{A^n R})$ of the state ρ. This expression, Proposition 4 in Section 5.2 and Lemma 12 in [1] imply (81).

Continuity bound (82) is obtained by using Proposition 6 in Section 6.1 twice and Corollary 2A in Section 3.

Continuity bound (83) is proved by noting that representation (19) implies

$$\bar{C}_{\mathrm{p}}(\Phi^{\otimes n}) = \sup_{\hat{\rho}} \left[I(B^n:C)_{\Phi^{\otimes n} \otimes \mathrm{Id}_C(\hat{\rho})} - I(E^n:C)_{\widehat{\Phi}^{\otimes n} \otimes \mathrm{Id}_C(\hat{\rho})} \right], \tag{85}$$

where the supremum is over all qc-states in $A^n C$. Inequality (83) is obtained by using Corollary 2A in Section 3, Proposition 4 in Section 5.2 twice and Lemma 12 in [1].

To prove the tightness of continuity bounds (79),(81) and (82) consider the family of erasure channels

$$\Phi_p(\rho) = \begin{bmatrix} (1-p)\rho & 0 \\ 0 & p\mathrm{Tr}\rho \end{bmatrix}, \quad p \in [0,1]. \tag{86}$$

from d-dimensional quantum system A to $(d+1)$-dimensional quantum system B. It is well known (see [6, 12]) that

$$C(\Phi_p) = C_\chi(\Phi_p) = (1-p)\log d, \tag{87}$$

$$Q(\Phi_p) = C_{\mathrm{p}}(\Phi_p) = \bar{C}_{\mathrm{p}}(\Phi_p) = \max\{(1-2p)\log d, 0\}. \tag{88}$$

By writing the channel Φ_p as the map $\rho \mapsto (1-p)\rho \oplus [p\mathrm{Tr}\rho]|\psi\rangle\langle\psi|$ from $\mathfrak{T}(\mathcal{H}_A)$ to $\mathfrak{T}(\mathcal{H}_A \oplus \mathcal{H}_\psi)$, where $\mathcal{H}_\psi$ is the space generated by $|\psi\rangle$, we see that the isometry

$$V_p : |\varphi\rangle \mapsto \sqrt{1-p}|\varphi\rangle \otimes |\psi\rangle \oplus \sqrt{p}|\psi\rangle \otimes |\varphi\rangle$$

from $\mathcal{H}_A$ into $\mathcal{H}_{BE}$, where $\mathcal{H}_E = \mathcal{H}_B = \mathcal{H}_A \oplus \mathcal{H}_\psi$, is a Stinespring isometry for Φ_p, i.e. $\Phi_p(\rho) = \mathrm{Tr}_E V_p \rho V_p^*$, for each p. Direct calculation shows that

$$\|V_{1/2-x} - V_{1/2}\| = \sqrt{2 - \sqrt{1-2x} - \sqrt{1+2x}} = x + o(x) \quad \text{as} \ \ x \to 0. \tag{89}$$

It follows from (87) and (88) that $C_*(\Phi_{1/2-x}) - C_*(\Phi_{1/2}) = x \log d$ for $C_* = C_\chi, C$ and that $C_*(\Phi_{1/2-x}) - C_*(\Phi_{1/2}) = 2x \log d$ for $C_* = Q, \bar{C}_{\mathrm{p}}, C_{\mathrm{p}}$.

Since (89) implies $\beta(\Phi_{1/2-x}, \Phi_{1/2}) \leq x + o(x)$ for small x, we see that continuity bounds (79),(81) and (82) are tight (for large d_A), while others are close-to-tight (up to factor 2 in the main term).

The last assertion of the theorem follows from the right inequality in (32) and monotonicity of the function $g(x)$. $\square$

8. Continuity bounds for basic capacities of energy-constrained infinite-dimensional channels.

In [9, 10] continuity bounds for classical and quantum capacities of energy-constrained infinite-dimensional channels with bounded energy amplification factor are obtained. In this section we use the results of Sections 5 and 6 to derive tight and close-to-tight continuity bounds for basic capacities valid for *arbitrary* energy-constrained infinite-dimensional channels.

When we consider transmission of information over infinite-dimensional quantum channels we have to impose constraints on states used for encoding information to be consistent with the physical implementation of the process. A typical physically motivated constraint is the boundedness of the average energy of states used for encoding information. For a single channel from A to B this constraint is expressed by the inequality

$$\mathrm{Tr}\, H_A \rho \leq E, \quad \rho \in \mathfrak{S}(\mathcal{H}_A), \tag{90}$$

where H_A is the Hamiltonian of the input system A, for n-copies of a channel it can be written as follows

$$\mathrm{Tr}\, H_{A^n} \rho \leq nE, \quad \rho \in \mathfrak{S}(\mathcal{H}_A^{\otimes n}), \tag{91}$$

where $H_{A^n} = H_A \otimes I_A \otimes \ldots \otimes I_A + \ldots + I_A \otimes \ldots \otimes I_A \otimes H_A$ is the Hamiltonian of the system A^n [6, 7, 8].

In this section we apply the results of Sections 5 and 6 to obtain continuity bounds for capacities of infinite-dimensional quantum channels with constraint (91) depending on E and the energy-constrained Bures distance between channels (introduced in Section 3) assuming that the Hamiltonian H_A satisfies condition (15). We consider all the basic capacities excepting the classical entanglement-assisted capacity C_{ea}, since the continuity bound for this capacity depending on the input energy bound and the energy-constrained diamond norm distance between quantum channels is obtained in [9, Proposition 7].

The Holevo capacity of any channel $\Phi : A \to B$ with constraint (91) is defined as:

$$C_\chi(\Phi, H_A, E) = \sup_{\mathrm{Tr}\, H_A \bar\rho \leq E} \chi(\{p_i, \Phi(\rho_i)\}), \tag{92}$$

where the supremum is over all discrete ensembles of states in $\mathfrak{S}(\mathcal{H}_A)$ with the average energy not exceeding E.

By the Holevo-Schumacher-Westmoreland theorem adapted for constrained infinite-dimensional channels (see [7, Proposition 3]) the classical capacity of any channel $\Phi : A \to B$ with constraint (91) is given by the regularized expression

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

Detailed analysis of the energy-constrained quantum and private capacities in the context of general-type infinite-dimensional channels[*] has been made recently by Wilde

and Qi in [8]. The results in [8] give considerable reasons to conjecture validity of the following generalizations of the Lloyd-Devetak-Shor theorem and of the Devetak theorem to energy-constrained infinite-dimensional channels:

- the quantum capacity of any channel $\Phi : A \to B$ with constraint (91) is given by the regularized expression

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

  where $\bar{Q}(\Phi, H_A, E)$ is the least upper bound of the coherent information $I_c(\Phi, \rho) \doteq I(\Phi, \rho) - H(\rho)$ on the set of all input states $\rho \in \mathfrak{S}(\mathcal{H}_A)$ satisfying (90).

- the private capacity of any channel $\Phi : A \to B$ with constraint (91) is given by the regularized expression

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

  where $\bar{C}_{\mathrm{p}}(\Phi, H_A, E) = \sup_{\mathrm{Tr} H_A \bar{\rho} \leq E} \left[\chi(\{p_i, \Phi(\rho_i)\}) - \chi(\{p_i, \widehat{\Phi}(\rho_i)\}) \right]$ (the supremum is over all discrete ensembles of states in $\mathfrak{S}(\mathcal{H}_A)$ with the average energy not exceeding E and $\widehat{\Phi}$ is the complementary channel to the channel Φ defined in (27)).

The uniform finite-dimensional approximation theorem for all the above capacities obtained in [33] and the analog of Theorem 1 for constrained channels with finite-dimensional input space show that all the functions

$$\Phi \mapsto C_*(\Phi, H_A, E), \quad C_* = C_\chi, \ C, \ \bar{Q}, \ Q, \ \bar{C}_{\mathrm{p}}, \ C_{\mathrm{p}},$$

are uniformly continuous on the set of all channels from A to arbitrary system B with respect to the strong (pointwise) convergence (34) provided that the Hamiltonian H_A satisfies condition (15). But the continuity bounds obtained in [33, Theorem 2] by this way are too rough and can not be used in applications. The following theorem presents more sharp and usable continuity bounds for all the above functions excepting $\bar{Q}(\Phi, H_A, E)$ which depend only on the input energy bound E and the energy-constrained Bures distance β_E between quantum channels defined in (35) (the tight continuity bound for the quantity $\bar{Q}(\Phi, H_A, E)$ depending on E coincides with the continuity bound for the capacity $C_{\mathrm{ea}}(\Phi, H_A, E)$ obtained in [9, Proposition 7]).

Theorem 2. *Let H_A be the Hamiltonian of a system A satisfying condition (15). Let $s = 0$ if the function $E \mapsto \bar{F}_{H_A}(E)/\sqrt{E}$ is non-increasing and $s = 1$ otherwise.♯ Then for arbitrary quantum channels Φ and Ψ from the system A to any system B the following inequalities hold*

$$|C_\chi(\Phi, H_A, E) - C_\chi(\Psi, H_A, E)| \leq T_{s,0}(E, \varepsilon) + g(\varepsilon) + 2\varepsilon \log 2, \tag{93}$$

$$|C(\Phi, H_A, E) - C(\Psi, H_A, E)| \leq T_{s,0}(E, \varepsilon) + g(\varepsilon) + 2\varepsilon \log 2, \tag{94}$$

$$|Q(\Phi, H_A, E) - Q(\Psi, H_A, E)| \leq T_{s,1}(E, \varepsilon) + g(\varepsilon) + 2\varepsilon \log 2, \tag{95}$$

♯ The function $\bar{F}_{H_A}$ is defined in (12).

$$|\bar{C}_{\mathrm{p}}(\Phi, H_A, E) - \bar{C}_{\mathrm{p}}(\Psi, H_A, E)| \leq 2T_{s,0}(E, \varepsilon) + 2g(\varepsilon) + 4\varepsilon \log 2, \qquad (96)$$

$$|C_{\mathrm{p}}(\Phi, H_A, E) - C_{\mathrm{p}}(\Psi, H_A, E)| \leq 2T_{s,1}(E, \varepsilon) + 2g(\varepsilon) + 4\varepsilon \log 2, \qquad (97)$$

where $\varepsilon = \beta_E(\Phi, \Psi)$ and $T_{s,t}(E, \varepsilon)$ is the quantity defined in Proposition 5. Inequalities (93)-(97) also hold with $\varepsilon = \sqrt{\|\Phi - \Psi\|_\diamond^E}$.

If A is the ℓ-mode quantum oscillator then inequalities (93)-(97) hold with $T_{s,t}(E, \varepsilon)$ replaced by the quantity $P_r(E, \varepsilon)$ defined in (65) for any $r \in (0, 1]$ such that $\varepsilon \leq 1/r$, where $\varepsilon = \beta_E(\Phi, \Psi)$ or $\varepsilon = \sqrt{\|\Phi - \Psi\|_\diamond^E}$. In this case continuity bounds (95) with optimal r is tight for large E, other continuity bounds are close-to-tight (up to factor 2 in the main term).

Proof. The continuity bound (93) for the Holevo capacity directly follows from its definition (92) and Proposition 8 in Section 6.2.

The continuity bound (94) is proved by using representation (19), Proposition 5 in Section 5.2 and Lemma 12 in [1]. The possibility to set $t = 0$ in (94) is due to the fact the supremum in the definition of $C_\chi(\Phi^{\otimes n}, H_{A^n}, nE)$ can be taken over all ensembles $\{p_i, \rho_i\}$ of states in $\mathfrak{S}(\mathcal{H}_A^{\otimes n})$ with the average state $\bar{\rho}$ such that $\mathrm{Tr} H_A \bar{\rho}_{A_k} \leq E$ for all $k = \overline{1, n}$. This can be easily shown by using the symmetry arguments (see the proof of Proposition 6 in [9]).

The continuity bound (95) is proved by using expression (84), Proposition 5 in Section 5.2 and Lemma 12 in [1].

The continuity bound (96) is obtained by using Proposition 8 in Section 6.2 twice and Corollary 2B in Section 3.

The continuity bound (97) is proved by using representation (19), Corollary 2B in Section 3, Proposition 5 in Section 5.2 twice and Lemma 12 in [1].

The assertion concerning the ℓ-mode quantum oscillator follows from Corollary 3 in Section 5.2. To show that the continuity bounds are tight or close-to-tight consider the family of erasure channels (86) from the ℓ-mode quantum oscillator to the system described by the space $\mathcal{H}_B = \mathcal{H}_A \oplus \{c\psi\}$. By generalizing the arguments from [6, Ch.10] it is easy to show that $C(\Phi_p, H_A, E) = C_\chi(\Phi_p, H_A, E) = (1 - p)M$ and

$$Q(\Phi_p, H_A, E) = C_{\mathrm{p}}(\Phi_p, H_A, E) = \bar{C}_{\mathrm{p}}(\Phi_p, H_A, E) = \max\{(1 - 2p)M, 0\},$$

where $M = F_{H_A}(E)$. It follows that $C_*(\Phi_{1/2-x}, H_A, E) - C_*(\Phi_{1/2}, H_A, E) = xM$, $C_* = C_\chi, C$ and that $C_*(\Phi_{1/2-x}, H_A, E) - C_*(\Phi_{1/2}, H_A, E) = 2xM$, $C_* = Q, \bar{C}_{\mathrm{p}}, C_{\mathrm{p}}$.

By the proof of Theorem 1 $\beta_E(\Phi_{1/2-x}, \Phi_{1/2}) \leq \beta(\Phi_{1/2-x}, \Phi_{1/2}) \leq x + o(x)$ for small x. So, by using Remark 2 in Section 5.2 it is easy to show that continuity bound (95) is tight for large E and others are close-to-tight (up to factor 2 in the main term). $\square$

Acknowledgments

for stimulating discussion. Special thanks to M.M.Wilde for valuable communication concerning capacities of infinite-dimensional channels with energy constraints.

Appendix: the addition to the proof of Proposition 1

Denote by $\beta'_E(\Phi, \Psi)$ the r.h.s. of (37). Let $\mathfrak{C}_{H_A,E}$ be the subset of $\mathfrak{S}(\mathcal{H}_A)$ determined by the inequality $\mathrm{Tr} H_A \rho \leq E$ and $\mathcal{N}(\Phi, \Psi) = \bigcup V_\Phi^* V_\Psi$, where the union is over *all* common Stinespring representations (31). Then it is easy to see that

$$\beta'_E(\Phi, \Psi) = \inf_{N \in \mathcal{N}(\Phi,\Psi)} \sup_{\rho \in \mathfrak{C}_{H_A,E}} \sqrt{2 - 2\Re \mathrm{Tr} N \rho} \tag{A.1}$$

Following the proof of Theorem 1 in [30] show that $\mathcal{N}(\Phi, \Psi)$ coincides with the set

$$\mathcal{M}(\Phi, \Psi) \doteq \{V_\Phi^*(I_B \otimes C)V_\Psi \,|\, C \in \mathfrak{B}(\mathcal{H}_E), \|C\| \leq 1\},$$

defined via some *fixed* common Stinespring representation (31). It would imply, in particular, that $\mathcal{M}(\Phi, \Psi)$ does not depend on this representation.

To show that $\mathcal{M}(\Phi, \Psi) \subseteq \mathcal{N}(\Phi, \Psi)$ it suffices to find for any contraction $C \in \mathfrak{B}(\mathcal{H}_E)$ a common Stinespring representation for Φ and Ψ with the isometries $\tilde{V}_\Phi$ and $\tilde{V}_\Psi$ from $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathcal{H}_{\tilde{E}}$ such that $\tilde{V}_\Phi^* \tilde{V}_\Psi = V_\Phi^*(I_B \otimes C)V_\Psi$.

Let $\mathcal{H}_{\tilde{E}} = \mathcal{H}_E^1 \oplus \mathcal{H}_E^2$, where $\mathcal{H}_E^1$ and $\mathcal{H}_E^2$ are copies of $\mathcal{H}_E$. For given C define the isometries $\tilde{V}_\Phi$ and $\tilde{V}_\Psi$ from $\mathcal{H}_A$ into $\mathcal{H}_B \otimes (\mathcal{H}_{E_1} \oplus \mathcal{H}_{E_2})$ by setting

$$\tilde{V}_\Phi|\varphi\rangle = V_\Phi|\varphi\rangle \oplus |0\rangle, \quad \tilde{V}_\Psi|\varphi\rangle = (I_B \otimes C)V_\Psi|\varphi\rangle \oplus \left(I_B \otimes \sqrt{I_E - C^*C}\right)V_\Psi|\varphi\rangle$$

for any $\varphi \in \mathcal{H}_A$, where we assume that the isometries V_Φ and V_Ψ act from $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathcal{H}_E^1$ and $\mathcal{H}_B \otimes \mathcal{H}_E^2$ correspondingly, while the contraction C acts from $\mathcal{H}_E^2$ to $\mathcal{H}_E^1$. It is easy to see that the isometries $\tilde{V}_\Phi$ and $\tilde{V}_\Psi$ form a common Stinespring representation for the channels Φ and Ψ with the required property.

To prove that $\mathcal{N}(\Phi, \Psi) \subseteq \mathcal{M}(\Phi, \Psi)$ take any common Stinespring representation for Φ and Ψ with the isometries $\tilde{V}_\Phi$ and $\tilde{V}_\Psi$ from $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathcal{H}_{\tilde{E}}$. By Theorem 6.2.2 in [6] there exist partial isometries W_Φ and W_Ψ from $\mathcal{H}_E$ to $\mathcal{H}_{\tilde{E}}$ such that $\tilde{V}_\Phi = (I_B \otimes W_\Phi)V_\Phi$ and $\tilde{V}_\Psi = (I_B \otimes W_\Psi)V_\Psi$. So, $\tilde{V}_\Phi^* \tilde{V}_\Psi = V_\Phi^*(I_B \otimes W_\Phi^* W_\Psi)V_\Psi \in \mathcal{M}(\Phi, \Psi)$, since $\|W_\Phi^* W_\Psi\| \leq 1$.

Since $\mathcal{N}(\Phi, \Psi) = \mathcal{M}(\Phi, \Psi)$, the infimum in (A.1) can be taken over the set $\mathcal{M}(\Phi, \Psi)$. This implies

$$\begin{aligned}\beta'_E(\Phi, \Psi) &= \inf_{C \in \mathfrak{B}_1(\mathcal{H}_E)} \sup_{\rho \in \mathfrak{C}(\mathcal{H}_A,E)} \sqrt{2 - 2\Re \mathrm{Tr} V_\Phi^*(I_B \otimes C)V_\Psi \rho} \\ &= \sup_{\rho \in \mathfrak{C}(\mathcal{H}_A,E)} \inf_{C \in \mathfrak{B}_1(\mathcal{H}_E)} \sqrt{2 - 2\Re \mathrm{Tr} V_\Phi^*(I_B \otimes C)V_\Psi \rho},\end{aligned} \tag{A.2}$$

where the possibility to change the order of the optimization follows from Ky Fan's minimax theorem [44] and the σ-weak compactness of the unit ball $\mathfrak{B}_1(\mathcal{H}_E)$ of $\mathfrak{B}(\mathcal{H}_E)$ [45]. It is easy to see that

$$\sup_{C \in \mathfrak{B}_1(\mathcal{H}_E)} \Re \mathrm{Tr} V_\Phi^*(I_B \otimes C)V_\Psi \rho = \sup_{C \in \mathfrak{B}_1(\mathcal{H}_E)} |\langle V_\Phi \otimes I_R \varphi | I_{BR} \otimes C | V_\Psi \otimes I_R \varphi\rangle|, \tag{A.3}$$

where R is any system and φ is a purification of ρ, i.e. a vector in $\mathcal{H}_{AR}$ such that $\mathrm{Tr}_R|\varphi\rangle\langle\varphi| = \rho$.

Since for *any* common Stinespring representation (31) and any system R the vectors $V_\Phi \otimes I_R|\varphi\rangle$ and $V_\Psi \otimes I_R|\varphi\rangle$ in $\mathcal{H}_{BER}$ are purifications of the states $\Phi \otimes \mathrm{Id}_R(|\varphi\rangle\langle\varphi|)$ and $\Psi \otimes \mathrm{Id}_R(|\varphi\rangle\langle\varphi|)$ in $\mathfrak{S}(\mathcal{H}_{BR})$, by using the relation $\mathcal{N}(\Phi, \Psi) = \mathcal{M}(\Phi, \Psi)$ proved before and Uhlmann's theorem [46][12, Ch.9] it is easy to show that the quantity in the r.h.s. of (A.3) coincides with the fidelity of the states $\Phi \otimes \mathrm{Id}_R(|\varphi\rangle\langle\varphi|)$ and $\Psi \otimes \mathrm{Id}_R(|\varphi\rangle\langle\varphi|)$. This and (A.2) implies that $\beta'_E(\Phi, \Psi) = \beta_E(\Phi, \Psi)$.

Assertion c) can be derived from the attainability of the infima in (A.2) which follows from the σ-weak compactness of the unit ball $\mathfrak{B}_1(\mathcal{H}_E)$.

References

[1] Leung D, Smith G 2009, *Commun. Math. Phys.* **292** 201-215.
[2] Alicki R, Fannes M 2004 *Journal of Physics A: Mathematical and General* **37**:5 L55-L57.
[3] Audenaert K M R 2007 *J. Math. Phys. A: Math. Theor.* **40**(28) 8127-8136.
[4] Fannes M 1973 *Commun. Math. Phys.* **31** 291-294.
[5] Winter A 2016 *Commun. Math. Phys.* **347**:1 291-313.
[6] Holevo A S 2012 *Quantum systems, channels, information. A mathematical introduction* (Berlin, DeGruyter).
[7] Holevo A S 2003 *Probability Theory and Applications* **48**:2 359-374.
[8] Wilde M M, Qi H 2016 Energy-constrained private and quantum capacities of quantum channels arXiv:1609.01997 [quant-ph].
[9] Shirokov M E 2018 *Problems of Information Transmission* **54**:1 20-33.
[10] Winter A 2017 Energy-Constrained Diamond Norm with Applications to the Uniform Continuity of Continuous Variable Channel Capacities arXiv:1712.10267 [quant-ph].
[11] Weedbrook C, Pirandola S, Garcia-Patron R, Cerf N J, Ralph T C, Shapiro J H, Lloyd S 2012 *Rev. Mod. Phys.* **84**, 621.
[12] Wilde M M 2013 *Quantum Information Theory* (Cambridge University Press).
[13] Lindblad G 1974 *Commun. Math. Phys.* **39** 111-119.
[14] Lindblad G 1973 *Commun. Math. Phys.* **33** 305-322.
[15] Li N, Luo S 2011 *Phys. Rev. A* **84**, 042124.
[16] Shirokov M E 2016 *Sbornik: Mathematics* **207**:5, 724-768.
[17] Shirokov M E 2017 *J. Math. Phys.* **58**:10, 102202, 29 pp.
[18] Wehrl A 1978 *Rev. Mod. Phys.* **50**, 221-250.
[19] Mosonyi M, Hiai F 2011 *IEEE Trans. Inf. Theory* **57**4, 2474-2487.
[20] Shirokov M E, 2018 *Reports on Mathematical Physics* **81**:1 81-104.
[21] Holevo A S 1973 *Problems of Information Transmission* **9** 177-183.
[22] Oreshkov O, Calsamiglia J 2009 *Phys. Rev. A* **79** 032336.
[23] Holevo A S, Shirokov M E 2005 *Theory of Probability and Applications* **50**:1 86-98.
[24] Bogachev V I 2007 *Measure theory* (Springer-Verlag Berlin Heidelberg).
[25] Parthasarathy K R 1967 *Probability measures on metric spaces* (Academic Press, New York and London).
[26] Aharonov D, Kitaev A, Nisan N 1998 *Proc. 30th STOC*, pp. 20-30, (ACM Press)
[27] Paulsen V I 2002 *Completely Bounded Maps and Operator Algebras* (Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge).
[28] Belavkin V P, D'Ariano G M, Raginsky M 2005 *J.Math.Phys.* **46** 062106.
[29] Kretschmann D, Schlingemann D, Werner R F. 2008 *IEEE Trans. Inform. Theory* **54**:4 1708-1717.
[30] Kretschmann D, Schlingemann D, Werner R F. 2008 *J. Funct. Anal.* **255**:8 1889-1904.

[31] Pirandola S, Laurenza R, Ottaviani C, Banchi L 2017 *Nat. Comm.* **8**, 15043.

[32] Holevo A S 2006 *Theory of Probability and Applications* **51** 133-143.

[33] Shirokov M E 2018 *Quantum Information Processing* doi: 10.1007/s11128-018-2070-z.

[34] Holevo A S 1998 *IEEE Trans. Inf. Theory* **44**:1 269-273.

[35] Schumacher B, Westmoreland M D 1997 *Phys. Rev. A* **56**, 131-138.

[36] King C 2002 *J. Math. Phys.* **43**:10 4641-4653.

[37] Amosov G G 2006 *Problems of Information Transmission* **42**:2 69-76.

[38] Bennett C H, Shor P W, Smolin J A, Thapliyal A V 1999 *Phys. Rev. Lett.* **83**, 3081-3084.

[39] Devetak I 2005 *IEEE Trans. Inf. Theory* **51**:1 44-55.

[40] Lloyd S 1997 *Physical Review A* **55**:3, 1613-1622.

[41] Cubitt T S, Ruskai M B, Smith G 2008 *J. Math. Phys.* **49** 102-104.

[42] Holevo A S, Werner R F 2001 *Physical Review A* **63**:3 032312.

[43] Wolf M M, Perez-Garcia D, Giedke G 2007 *Physical Review Letters* **98**:13 130501.

[44] Simons S 1998 *Minimax and Monotonicity* (Springer, Berlin).

[45] Bratteli O, Robinson D W 1979 *Operators algebras and quantum statistical mechanics*, vol.I, (Springer Verlag, New York-Heidelberg-Berlin).

[46] Uhlmann A 1976 *Rep. Math. Phys.* **9** 273.